

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2005, 03:18:11 ; Search time 5462 Seconds
(without alignments)

11041.908 Million cell updates/sec

Title: US-10-826-929A-1

Perfect score: 1061

Sequence: 1 agcaaaagcaggggatattt.....taccagaaaagcaaatcaga 1061

Scoring table: IDENTITY NUC

Gapop 10⁻⁰, Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*

2: gb_in.*

3: gb_env.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pr.*

9: gb_ro.*

10: gb_sts.*

11: gb_sy.*

12: gb_un.*

13: gb_vi.*

14: gb_htg.*

15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1061	100.0	1061	13	AF197241 Influenza
2	1056.2	99.5	1061	13	AY273168 Influenza
3	1053	99.2	1061	13	AF197248 Influenza
4	1051.4	99.1	1061	13	AF197247 Influenza
5	1048.2	98.8	1762	13	FLAH33C
6	1043.4	98.3	1060	13	EIVY14059
7	1041.8	98.2	1060	13	EIVY14060
8	1041.8	98.2	1762	13	FLAH33D
9	1039.2	97.9	1100	13	EIVY14058
10	1038.6	97.9	1762	13	FLAH33F
11	1035.4	97.6	1061	13	AF197242
12	1033.8	97.4	1762	6	BD244631
13	1033.8	97.4	1762	6	AR254631
14	1033.8	97.4	1762	6	AR343239
15	1033.8	97.4	1762	6	AR455506
16	1033.8	97.4	1762	6	AR609062
17	1033.8	97.4	1762	13	FLAH33A
18	1032.2	97.3	1762	6	BD244629

19	1032.2	97.3	1762	6	AR254629	AR254629 Sequence
20	1032.2	97.3	1762	6	AR343237	AR343237 Sequence
21	1032.2	97.3	1762	6	AR455504	AR455504 Sequence
22	1032.2	97.3	1762	6	AR609060	AR609060 Sequence
23	1030.6	97.1	1061	13	AF197245	AF197245 Influenza
24	1030.6	97.1	1762	13	FLAH33E	L39917 Influenza A
25	1029	97.0	1061	13	AF197249	AF197249 Influenza A
26	1027.4	96.8	1762	13	FLAH33C	L39915 Influenza A
27	1022.6	96.4	1061	13	AF197243	AF197243 Influenza A
28	1022.6	96.4	1061	13	AF197244	AF197244 Influenza A
29	1022.6	96.4	1090	13	EIVY14056	Y14056 Influenza A
30	1022.6	96.4	1762	13	IVAH691	X95637 Influenza A
31	1021.2	96.2	1762	13	IVAH991	X95638 Influenza A
32	1019.2	96.1	1061	13	AY273167	AY273167 Influenza A
33	1019.2	96.1	1032	13	IVTAN931	X85088 Influenza A
34	1019.2	96.1	1698	6	AR490205	AR490205 Sequence
35	1019.2	96.1	1698	6	AX018718	AX018718 Sequence
36	1018.4	96.0	1762	6	AR011427	AR011427 Sequence
37	1018.4	96.0	1762	6	T18065	T18065 Sequence 30
38	1016.8	95.8	1762	13	FLAET85HA	M24726 Influenza A
39	1016.2	95.8	1762	13	EIVH3A	L27597 Influenza A
40	1015.6	95.7	1040	13	FLAH33NHAD	D30680 Influenza A
41	1014.6	95.6	1061	13	AF197246	AF197246 Influenza A
42	1014.6	95.6	1762	13	AY855341	AY855341 Influenza A
43	1014.2	95.6	1099	13	FLAH33NHAD	D30686 Influenza A
44	1011.8	95.4	1093	13	FLAH33NHAE	D30681 Influenza A
45	1010.4	95.2	1762	13	FLAEK87HA	M24728 Influenza A

ALIGNMENTS

RESULT 1

AF197241

LOCUS

DEFINITION

AF197241

ACCESSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

/translation="MKTIIILLTHWVSONPTSGNNTATILCLGHHAIVANGTLVKTI
TDOIEVNTATELVQSIGKICNNYSKVLDRGNTCLIDAMLGDPDCHDFVQYENWDLF
IERSSAFSCNCPYDIPDYASLSIVASSGTLFTAGFTWTGVTQNGRSGACKRSAD
SFVSRNLWUTKSGNSYPTLNVTPNNKNFDKLYIWIHHPSSNQOQTELYIIOESGRVT
VSTKRSQOITVFNIGSRPNVRGQSGRSIYWTIVKPGDILMINSNGLVAPRGYFKLK
TGKSSVMRSDAPIDICVSECITPNGSI PNDKPFQNVKVTYKCPKVIQNTLKLATG
MKNIPKQIR"
30..74
/gene="HA1"
75..>1061
/gene="HA1"
/product="hemagglutinin"

sig_peptide

mat_peptide

ORIGIN

Query Match 100.0%; Score 1061; DB 13; Length 1061;

Best Local Similarity 100.0%; Pred. No. 1.5e-235;

Matches 1061; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	AGCAAAAGCAGGGGATATTCTCTCAATCATGAAGACAACCATTTATTTGATACTACTGA	60
DB	1	AGCAAAAGCAGGGGATATTCTCTCAATCATGAAGACAACCATTTATTTGATACTACTGA	60
QY	61	CCCATTTGGGTCTACAGTCAAAACCCAAACCCAGTGGAAACACACAGCCACATTATGCTCG	120
DB	61	CCCATTTGGGTCTACAGTCAAAACCCAAACCCAGTGGAAACACACAGCCACATTATGCTCG	120
QY	121	GACACCATGCGTAGCAATGGACATTTGGTAAACCAATTAAGTGAACCAATTTGAGG	180
DB	121	GACACCATGCGTAGCAATGGACATTTGGTAAACCAATTAAGTGAACCAATTTGAGG	180
QY	181	TGCAAAATGCTACTGAATAGTTTCAGAGCATTTCAATAGGAAATATGCAACCACTCAT	240
DB	181	TGCAAAATGCTACTGAATAGTTTCAGAGCATTTCAATAGGAAATATGCAACCACTCAT	240
QY	241	ATAAAGTCTTAGATGAAGAAATTCACATTAATAGATGCAATGCTAGGAGACCCCACT	300
DB	241	ATAAAGTCTTAGATGAAGAAATTCACATTAATAGATGCAATGCTAGGAGACCCCACT	300
QY	301	GTGATCTCTCCAGTATGAAATGGACCTCTTCATAGAAAGCAGCGCTTTCAGCA	360
DB	301	GTGATCTCTCCAGTATGAAATGGACCTCTTCATAGAAAGCAGCGCTTTCAGCA	360
QY	361	ATTGCTACCCATATGACATCCCTGACTATGCAATCGCTCCGGTCCATTTAGCATCCTCAG	420
DB	361	ATTGCTACCCATATGACATCCCTGACTATGCAATCGCTCCGGTCCATTTAGCATCCTCAG	420
QY	421	GAACTATTAGAAATTCACAGCAGAGGATTCATGAGGACAGGTGTCACTCAAAAACGGAAGAA	480
DB	421	GAACTATTAGAAATTCACAGCAGAGGATTCATGAGGACAGGTGTCACTCAAAAACGGAAGAA	480
QY	481	GTGAGCGCTGCAAAAGGGATCAGCGATAGTTTCTTTAGCCGACCTGAATTTGGCTAACAA	540
DB	481	GTGAGCGCTGCAAAAGGGATCAGCGATAGTTTCTTTAGCCGACCTGAATTTGGCTAACAA	540
QY	541	AATCTGGAATCTTACCCACATTTGAATGTGACATTCCTTAACAATAAAAAATTTTCGACA	600
DB	541	AATCTGGAATCTTACCCACATTTGAATGTGACATTCCTTAACAATAAAAAATTTTCGACA	600
QY	601	AACATATACATCTGGGGATTCATCACCCGAGCTCAAAACCAACGACGACAGAAATTTGACAA	660
DB	601	AACATATACATCTGGGGATTCATCACCCGAGCTCAAAACCAACGACGACAGAAATTTGACAA	660
QY	661	TCCAAGATCAGCAGCAGTAACTCTCAACAAAAGAGCTCAACAAACGATAGTCCCTTA	720
DB	661	TCCAAGATCAGCAGCAGTAACTCTCAACAAAAGAGCTCAACAAACGATAGTCCCTTA	720
QY	721	ATATCGGATCTAGACCGTGGGTTAGGGGTCAATCAGGAGGATGAAGCATATCTGGACCA	780
DB	721	ATATCGGATCTAGACCGTGGGTTAGGGGTCAATCAGGAGGATGAAGCATATCTGGACCA	780
QY	781	TTGTAAACCTGAGATATCTTAATGATAACAGTAAATGCGCACTTAGTTGCAACCGGGG	840
DB	781	TTGTAAACCTGAGATATCTTAATGATAACAGTAAATGCGCACTTAGTTGCAACCGGGG	840

QY	841	GATATTTTAAATTTGAAACAGGAGAAAGCTCTGTATGATGAGATCAGATGCACCCATAGACA	900
DB	841	GATATTTTAAATTTGAAACAGGAGAAAGCTCTGTATGATGAGATCAGATGCACCCATAGACA	900
QY	901	TTTGTGTGTCTGAATGTATTACCAAAATGGAAGCATCCCAACGACAAACCATTTTCAAA	960
DB	901	TTTGTGTGTCTGAATGTATTACCAAAATGGAAGCATCCCAACGACAAACCATTTTCAAA	960
QY	961	ATGTGAACAAAGTTACATATGGAATAATGCCCAAGTATATCAGGCAAAAACATTTAAAGC	1020
DB	961	ATGTGAACAAAGTTACATATGGAATAATGCCCAAGTATATCAGGCAAAAACATTTAAAGC	1020
QY	1021	TGGCCACTGGGATGAGGAATATACCAAGAAACCAATCAGA	1061
DB	1021	TGGCCACTGGGATGAGGAATATACCAAGAAACCAATCAGA	1061

RESULT 2

AY273168

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..1061

/organism="Influenza A virus

/db_xref="taxon:241186"

/mol_type="genomic RNA"

/isolates="A/Eg/Oklahoma/2000"

/db_xref="taxon:241186"

30..>1061

/gene="HA1"

30..>1061

/gene="HA1"

/codon_start=1

/product="hemagglutinin precursor"

/protein_id="AAQ18435.1"

/db_xref="GI:33415852"

/translation="MKTIIILLTHWVSONPTSGNNTATILCLGHHAIVANGTLVKTI

TDOIEVNTATELVQSIGKICNNYSKVLDRGNTCLIDAMLGDPDCHDFVQYENWDLF

IERSSAFSCNCPYDIPDYASLSIVASSGTLFTAGFTWTGVTQNGRSGACKRSAD

SFVSRNLWUTKSGNSYPTLNVTPNNKNFDKLYIWIHHPSSNQOQTELYIIOESGRVT

VSTKRSQOITVFNIGSRPNVRGQSGRSIYWTIVKPGDILMINSNGLVAPRGYFKLK

TGKSSVMRSDAPIDICVSECITPNGSI PNDKPFQNVKVTYKCPKVIQNTLKLATG

MKNIPKQIR"

ORIGIN

Query Match

Best Local Similarity

Matches 1058; Conservative

99.5%; Score 1056.2; DB 13; Length 1061;

99.7%; Pred. No. 1.9e-234;

0; Mismatches 3; Indels 0; Gaps 0;

QY

QY	61	CCCATTTGGGCTCTACAGTCAAAACCCCAACCAAGTGGAAACAACACAGCCACATTTATGTCGG	120
Db	61	CCCATTTGGGCTCTACAGTCAAAACCCCAACCAAGTGGAAACAACACAGCCACATTTATGTCGG	120
QY	121	GACACCATGCACTAGCAAAATGGAAATGGTGTAAAAACAATAAATCACTGATGACCAAAATTCAGG	180
Db	121	GACACCATGCACTAGCAAAATGGAAATGGTGTAAAAACAATAAATCACTGATGACCAAAATTCAGG	180
QY	181	TGACAAATGCTACTGAATTTAGTTCAGAGCAATTTCAATAGGGAATAATGCAACAACTCAT	240
Db	181	TGACAAATGCTACTGAATTTAGTTCAGAGCAATTTCAATAGGGAATAATGCAACAACTCAT	240
QY	241	ATAAGTCTAGATGGAAGAAATTTGCAATTAATAGATGCAATGCTAGGAGACCCCACT	300
Db	241	ATAAGTCTAGATGGAAGAAATTTGCAATTAATAGATGCAATGCTAGGAGACCCCACT	300
QY	301	GTGATGCTTCCAGTATGAGAAATTTGGGACCTCTTCATAGAAAGAGCAGCGCTTTCAGCA	360
Db	301	GTGATGCTTTCAGTATGAGAAATTTGGGACCTCTTCATAGAAAGAGCAGCGCTTTCAGCA	360
QY	361	ATTGCTACCCATATGACATCCCTGACATATGATCGCTCCGGTCCATTTGATGATCCTCAG	420
Db	361	ATTGCTACCCATATGACATCCCTGACATATGATCGCTCCGGTCCATTTGATGATCCTCAG	420
QY	421	GAACTATTAGAAATTCACAGCAGAGGGATTCATGGAACAGGTGTCACATCAAAAACGGAAGAA	480
Db	421	GAACTATTAGAAATTCACAGCAGAGGGATTCATGGAACAGGTGTCACATCAAAAACGGAAGAA	480
QY	481	GTGAGGCTGCAAAAGGGGATCAGCCGATAGTTCTTTAGCCGACTGAAATTTGGCTAACAA	540
Db	481	GTGAGGCTGCAAAAGGGGATCAGCCGATAGTTCTTTAGCCGACTGAAATTTGGCTAACAA	540
QY	541	AATCTGAAACTCTTACCCACATTTGAATGTGACATGCTCTAACAAATAAAAAATTTGACA	600
Db	541	AATCTGAAACTCTTACCCACATTTGAATGTGACATGCTCTAACAAATAAAAAATTTGACA	600
QY	601	AATCTATACATCTGGGGATTCATCACCCGAGCTCAAAACCAACGACAGACAGAAATTTGACA	660
Db	601	AATCTATACATCTGGGGATTCATCACCCGAGCTCAAAACCAACGACAGACAGAAATTTGACA	660
QY	661	TCCAAGATCAGGACGATACAGTCTCAACAAAGAAAGTCAACAAACATAGTCGCTTA	720
Db	661	TCCAAGATCAGGACGATACAGTCTCAACAAAGAAAGTCAACAAACATAGTCGCTTA	720
QY	721	ATATCGGATCTAGACCGTGGTATAGGGGTCAATCAGGACGATAAGCATATCTGGACCA	780
Db	721	ATATCGGATCTAGACCGTGGTATAGGGGTCAATCAGGACGATAAGCATATCTGGACCA	780
QY	781	TTGTAAACCTGGAGATATCCTAATGATAACAGTAAATGGCAACTTAGTTGACCCGGGG	840
Db	781	TTGTAAACCTGGAGATATCCTAATGATAACAGTAAATGGCAACTTAGTTGACCCGGGG	840
QY	841	GATATTTAAATTTGAAACAGGGGAAAGCTCTGTATGATGATGATGACACCCATAGACA	900
Db	841	GATATTTAAATTTGAAACAGGGGAAAGCTCTGTATGATGATGATGACACCCATAGACA	900
QY	901	TTTGTGTGCTGATGATTTACACCAATGGAAGCATCCCAACGCAACACATTTTCAA	960
Db	901	TTTGTGTGCTGATGATTTACACCAATGGAAGCATCCCAACGCAACACATTTTCAA	960
QY	961	ATGTGAACAAAGTTTACATATGGAATAATGCCCCCAAGTATATCAGGCAAAAACATTTAAAGC	1020
Db	961	ATGTGAACAAAGTTTACATATGGAATAATGCCCCCAAGTATATCAGGCAAAAACATTTAAAGC	1020
QY	1021	TGGCCACTGGGATGAGGAATATACGAGAAAGCAAAATTCAGA	1061
Db	1021	TGGCCACTGGGATGAGGAATATACGAGAAAGCAAAATTCAGA	1061
RESULT 3			
AF197248		1061 bp	mRNA
LOCUS			VRL 08-JAN-2003


```
QY 481 GTGGAGCCTGCAGAAAGGGATCAGCGATAGTCTTTTTCAGCCGACTGAATTCGCTACAA 540
Db 481 GTGGAGCCTGCAGAAAGGGATCAGCGATAGTCTTTTTCAGCCGACTGAATTCGCTACAA 540
QY 541 AATCTGAAACTCTTACCCACATTTGAATGTGACAAATGCCTAACAAATAAAATTCGACA 600
Db 541 AATCTGAAACTCTTACCCACATTTGAATGTGACAAATGCCTAACAAATAAAATTCGACA 600
QY 601 AACTATACATCTGGGGATTTCATCACCAGGCTCAACCAACCAACAGCAGCAGAGATTTGACA 660
Db 601 AACTATACATCTGGGGATTTCATCACCAGGCTCAACCAACCAACAGCAGCAGAGATTTGACA 660
QY 661 TCCAAGAAATCAGGACGAGTAACTCTCAACCAAAAGAGTCAACAAACGATAGTCCCTA 720
Db 661 TCCAAGAAATCAGGACGAGTAACTCTCAACCAAAAGAGTCAACAAACGATAGTCCCTA 720
QY 721 ATATCGGATCTAGACCGTGGGTAGGGTCAATCAGCAGGATGAAGCATATCTGGACCA 780
Db 721 ATATCGGATCTAGACCGTGGGTAGGGTCAATCAGCAGGATGAAGCATATCTGGACCA 780
QY 781 TTGTAAACCTGGAGATCTCTAATGATAACAGTAAATGGAATTTAGTTGACCCGGG 840
Db 781 TTGTAAACCTGGAGATCTCTAATGATAACAGTAAATGGAATTTAGTTGACCCGGG 840
QY 841 GATATTTTAAATGAAACAGGAAAGCTCTGTAATGAGATCAGATGACCCCATAGACA 900
Db 841 GATATTTTAAATGAAACAGGAAAGCTCTGTAATGAGATCAGATGACCCCATAGACA 900
QY 901 TTTGTGTCTGATGATTAACCAAAATGGAAGCATCCCAACGACAAACCAATTTCAA 960
Db 901 TTTGTGTCTGATGATTAACCAAAATGGAAGCATCCCAACGACAAACCAATTTCAA 960
QY 961 ATGTGAACAAATGATGATGAAATGCCCCAGTATATCAGGCAAAAACATTTAAAGC 1020
Db 961 ATGTGAACAAATGATGATGAAATGCCCCAGTATATCAGGCAAAAACATTTAAAGC 1020
QY 1021 TGGCCACTGGGATGAGGAATATACCAAGAAAGCAATTCAGA 1061
Db 1021 TGGCCACTGGGATGAGGAATATACCAAGAAAGCAATTCAGA 1061
```

```
RESULT 5
FLAH3B 1762 bp ss-RNA linear VRL 01-NOV-2004
LOCUS Influenza A virus (A/equine/kentucky/1/94 (H3N8)) hemagglutinin
DEFINITION precursor (HA) gene, complete cds.
ACCESSION L39914
VERSION L39914.1 GI:722403
KEYWORDS Influenza A virus (A/equine/Kentucky/1/94 (H3N8))
SOURCE Influenza A virus (A/equine/Kentucky/1/94 (H3N8))
ORGANISM Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
Influenzavirus A.
REFERENCE 1 (bases 1 to 1762)
AUTHORS Daly J.M., Lai, A.C., Binns, M.M., Chambers, T.M., Barrandeguy, M. and
Mumford, J.A.
TITLE Antigenic and genetic evolution of equine H3N8 influenza A viruses
JOURNAL J. Gen. Virol. 77 (pt 4), 661-671 (1996)
PUBMED 8627254
FEATURES
source Location/Qualifiers
1..1762
/organism="Influenza A virus
(A/equine/kentucky/1/94 (H3N8))"
/mol_type="genomic RNA"
/strain="A/equine/Kentucky/1/94"
/serotype="H3N8"
/db_xref="taxon:39129"
/clone="R'AA"
/note="Equine-2"
30..1727
/genes="HA"
30..1727
CDS
```

```
/genes="HA"
/codon_start=1
/product="hemagglutinin precursor"
/protein_id="AAB36976.1"
/db_xref="GI:722404"
/translations="MKTIIILILTHWVYSONPTSGNNTATLCIGHHAVANGTLVKTI
TDQIEVTNATELVQISIGIKCNNSVRVLDRNCTLIDAMLDGPHCDVFOYENNDLF
IERSSAFSNCPYDI PDYASLRISIVASSGTLEAFSTGFTWTGTQNGRSACKRGSD
SFSRLNLTKSNSPTLVNTPNNKDFKLYIWIHHPSSNQOQTSLYIOESGRVT
VSTRSQQTIIPIGSRVWGSGRISIIYTIIVKPGDILMINSNGLVAPRGYFKLK
TGKSSVMRSDADIDICVSECITPNSIPNDKPFQNVNKVYTKCPKIRQNTLKLATG
MRNVPKQIRGIFGAIAGTENGWGMVGYFRYONSGTGOAADLKQTQAAIDQI
NKLNVIERTEKPHQIEKEFSEVEGRIODLEKYVEDTIDILMSYNAELLVALENOH
TIDLTAEWNKLFPEKTRROLRENAEDMGSCFKIYHKCDNACTGSRNGYVDHIYRD
EALNRFQIKGVLEKSGYKDWILWISFALSCFLICVLLGFIIMWACQGNIRNCICI"
30..74
sig_peptide
/genes="HA"
/notes="putative"
75..1061
mat_peptide
/genes="HA"
/product="HA1"
1065..1724
mat_peptide
/genes="HA"
/product="HA2"
ORIGIN
Query Match 98.8%; Score 1048.2; DB 13; Length 1762;
Best Local Similarity 99.2%; Pred. No. 1.2e-232;
Matches 1053; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 AGCAAAAGCAGGGGATATTTCTGCAATCATGAAGACCAACCATTTTTCATCTACTGA 60
Db 1 AGCAAAAGCAGGGGATATTTCTGCAATCATGAAGACCAACCATTTTTCATCTACTGA 60
QY 61 CCCATTGGGTCTACAGTCAAAACCCCAACCAAGTGGAAACACACAGCCACATTTGTCGG 120
Db 61 CCCATTGGGTCTACAGTCAAAACCCCAACCAAGTGGAAACACACAGCCACATTTGTCGG 120
QY 121 GACACCATGAGTAGCAAAATGGAACATTTGGTAAAAACAATACTGATGACAAATTCGAGG 180
Db 121 GACACCATGAGTAGCAAAATGGAACATTTGGTAAAAACAATACTGATGACAAATTCGAGG 180
QY 181 TGACAAATGCTACTGATTAATGTTTCAGAGCATTTTCAATAGGGAATAATGCAACACTCAT 240
Db 181 TGACAAATGCTACTGATTAATGTTTCAGAGCATTTTCAATAGGGAATAATGCAACACTCAT 240
QY 241 ATAAAGTTCTAGATGGAAGAAATTCACATTAATAGATGCAATTCGATGAGAGACCCCACT 300
Db 241 ATAGAGTTCTAGATGGAAGAAATTCACATTAATAGATGCAATTCGATGAGAGACCCCACT 300
QY 301 GTGATGCTTCAGATGAGAAATTTGGACCTCTTCATAGAAAGACAGCGCTTTTCAGCA 360
Db 301 GTGATGCTTCAGATGAGAAATTTGGACCTCTTCATAGAAAGACAGCGCTTTTCAGCA 360
QY 361 ATTGCTACCATATGATCCCTGATGATGATGCTCGCTCCGCTCCATTTGATAGCATCTCTCAG 420
Db 361 ATTGCTACCATATGATCCCTGATGATGATGCTCGCTCCGCTCCATTTGATAGCATCTCTCAG 420
QY 421 GAACATTAGAAATTCACAGCAGAGGGATTTCACATGAGCAGAGGTGTCACTCAAAAACGGAAGAA 480
Db 421 GAACATTAGAAATTCACAGCAGAGGGATTTCACATGAGCAGAGGTGTCACTCAAAAACGGAAGAA 480
QY 481 GTGAGCCTGCAGAAAGGGGATCAGCGATAGTTTCTTTTAGCCGACTGAATTCGCTAACAA 540
Db 481 GTGAGCCTGCAGAAAGGGGATCAGCGATAGTTTCTTTTAGCCGACTGAATTCGCTAACAA 540
QY 541 AATCTGAAACTCTTACCCACATTTGAATGTGACAAATGCCTAACAAATAAAATTCGACA 600
Db 541 AATCTGAAACTCTTACCCACATTTGAATGTGACAAATGCCTAACAAATAAAATTCGACA 600
QY 601 AACTATACATCTGGGGATTTCATCACCAGGCTCAACCAACCAACAGCAGCAGAGATTTGACA 660
Db 601 AACTATACATCTGGGGATTTCATCACCAGGCTCAACCAACCAACAGCAGCAGAGATTTGACA 660
```

661 TCCAGAAATCAGGACGAGTAACAGTCTCAACAAAAAGAGTCAACAAACGATAGTCCCTTA 720
Db
661 TCCAGAAATCAGGACGAGTAACAGTCTCAACAAAAAGAGTCAACAAACGATAGTCCCTTA 720
Qy 721 ATATCGGATCTAGACCGTGGGTAGGGTCAATCAGGACGAGTAAGCATATATCTGACCA 780
Db 721 ATATCGGATCTAGACCGTGGGTAGGGTCAATCAGGACGAGTAAGCATATATCTGACCA 780
Qy 781 TTGTAAAACTGAGATATCTTAATGATTAACAGTAAATGGCACTTGTGACCGGG 840
Db 781 TTGTAAAACTGAGATATCTTAATGATTAACAGTAAATGGCACTTGTGACCGGG 840
Qy 841 GATATTTTAAATGAAAAACAGGAAAAAGCTCTGTAATGAGATCAGATGCAACCCATAGACA 900
Db 841 GATATTTTAAATGAAAAACAGGAAAAAGCTCTGTAATGAGATCAGATGCAACCCATAGACA 900
Qy 901 TTGTGTGTCTGAATGTATTACACCAATGGAAGCATCCCCAACGACAAACCAATTTCAAA 960
Db 901 TTGTGTGTCTGAATGTATTACACCAATGGAAGCATCCCCAACGACAAACCAATTTCAAA 960
Qy 961 ATGTGAACAAAGTACATATGGAATGCCCAAGTATATCAGGCAAAACATTTTAAAGC 1020
Db 961 ATGTGAACAAAGTACATATGGAATGCCCAAGTATATCAGGCAAAACATTTTAAAGC 1020
Qy 1021 TGGCCACTGGGATGAGGAATATACAGAAAAAGCAATCAGA 1061
Db 1021 TGGCCACTGGGATGAGGAATGTAACAGAAAAAGCAATCAGA 1061

RESULT 6
EIVY14059 1060 bp RNA linear VRL 18-APR-2005
LOCUS
DEFINITION
Influenza A virus (A/equi 2/Alvdalen/96 (H3N8)) gene for
hemagglutinin H1 subunit, isolate A/equi 2/Alvdalen/96.
Y14059
ACCESSION
VERSION
Y14059.1 GI:2765371
KEYWORDS
hemagglutinin H1 subunit.
Influenza A virus (A/equi 2/Alvdalen/96 (H3N8))
Influenza A virus (A/equi 2/Alvdalen/96 (H3N8))
SOURCE
ORGANISM
Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
Influenzavirus A.
REFERENCE
1 (bases 1 to 1060)
Oxburgh, L., Akerblom, L., Fridberger, T., Klingeborn, B. and Linne, T.
AUTHORS
IDENTIFICATION OF TWO ANTIGENICALLY AND GENETICALLY DISTINCT
TITLES
LINEAGES OF H3N8 EQUINE INFLUENZA VIRUS IN SWEDEN
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 1060)
Oxburgh, L.H.
AUTHORS
Direct Submission
TITLES
Submitted (25-JUN-1997) Oxburgh L.H., Veterinary Microbiology,
JOURNAL
Swedish University of Agricultural Sciences, Biomedical Centre Box
585, s-751 23 Uppsala, SWEDEN

FEATURES
source
1. .1060
/organism="Influenza A virus (A/equi 2/Alvdalen/96 (H3N8))"
/location/Qualifiers
/mol_type="genomic RNA"
/serotype="H3N8"
/isolate="A/equi 2/Alvdalen/96"
/db_xref="taxon:217820"
<22> .>1060
/codon_start=1
/product="hemagglutinin H1 subunit"
/protein_id="CAA74387.1"
/db_xref="GI:2765372"
/db_xref="GOA:O56857"
/db_xref="InterPro:IPR000149"
/db_xref="InterPro:IPR001364"
/db_xref="UniProt/TrEMBL:O56857"
/translations="MKTIILILLTHWVYSONPTSGNNTATLCIGHHAVANGFLVKTII
TNDQIEVTNATLQVSIISIGIKCNNSVRLDGRNCTLIIDAMLDGDPHCPQYENWDLF

HERSAFNCYPIPIPDYASLRISIVASSGTLEFTEAGFTWTGVTQNGRSGACKRGSD
SFFSLNLTSGNSYPTLNVTPNNKNFDKLYIWGIHHPNSQQOTEIYIOESGRVT
VTKRSOOTIIPNIGSRPWRGOSGRISYWTIVKPGDILMINSNGLVAPGYFKLK
TKGSVMSSDAPIDICASECITPNISIPNDKFPQNVNKVYTGKPKYIQNTLKLATG
MRNVPEKQIRGI"
22. .1060
/number=1
22. .67
exon
sig_peptide
ORIGIN
Query Match 98.3%; Score 1043.4; DB 13; Length 1060;
Best Local Similarity 99.4%; Pred. No. 1.8e-231;
Matches 1047; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 9 CAGGGGATATTTCTGTCAATCATGAGACAAACATTATTTTGATCTACTGACCCCATTTG 68
Db 1 CAGGGGATATTTCTGTCAATCATGAGACAAACATTATTTTGATCTACTGACCCCATTTG 60
Qy 69 GTCTACAGTCAAAACCCAAACCCAGTGGAAACAAACAGCCACATTATGTCTGGGACACCAT 128
Db 61 GTCTACAGTCAAAACCCAAACCCAGTGGAAACAAACAGCCACATTATGTCTGGGACACCAT 120
Qy 129 GCAGTAGCAAAATGGAAACATTGGTAAAAACAATAAATGATGACCAAAATGAGGTGACAAAT 188
Db 121 GCAGTAGCAAAATGGAAACATTGGTAAAAACAATAAATGATGACCAAAATGAGGTGACAAAT 180
Qy 189 GCTACTGAATTTAGTTTCAGAGCATTTCAATAGGAAAAATATGCAACAACTCATATAAGTT 248
Db 181 GCTACTGAATTTAGTTTCAGAGCATTTCAATAGGAAAAATATGCAACAACTCATATAAGTT 240
Qy 249 CTAGATGGAAGAAATTTGCAATTATAGATGCAATGTAGGAGACCCCACTGTGTATGTC 308
Db 241 CTAGATGGAAGAAATTTGCAATTATAGATGCAATGTAGGAGACCCCACTGTGTATGTC 300
Qy 309 TTCCAGTATGAGAAATTTGGGACCTCTTCATAGAAAGAGCAGCGCTTTTCAGCAATTTCTTAC 368
Db 301 TTCCAGTATGAGAAATTTGGGACCTCTTCATAGAAAGAGCAGCGCTTTTCAGCAATTTCTTAC 360
Qy 369 CCATATGACATCCCTGACTATGTCATCGCTCCGTCCTTGTAGCATCTCTCAGGAAACATT 428
Db 361 CCATATGACATCCCTGACTATGTCATCGCTCCGTCCTTGTAGCATCTCTCAGGAAACATT 420
Qy 429 GAATTCACAGCAGAGGGATTTCATAGGACAGGTGTCACTCAAAAACGGAAGAGTGGAGCC 488
Db 421 GAATTCACAGCAGAGGGATTTCATAGGACAGGTGTCACTCAAAAACGGAAGAGTGGAGCC 480
Qy 489 TGCAAAAGGGGATCAGCCGATAGTTTCTTTAGCCGACTGAATTTGGCTAACAAATCTGGA 548
Db 481 TGCAAAAGGGGATCAGCCGATAGTTTCTTTAGCCGACTGAATTTGGCTAACAAATCTGGA 540
Qy 549 AACTCTTTACCCCACTTGAATGTGACAAATGCCTTAACAATAAATAATTTTCGACAACTATAC 608
Db 541 AACTCTTTACCCCACTTGAATGTGACAAATGCCTTAACAATAAATAATTTTCGACAACTATAC 600
Qy 609 ATCTGGGGATTCATCACCCGAGCTCAAAACCAACAGCAGACAGAAATTTGTACATCCAGAA 668
Db 601 ATCTGGGGATTCATCACCCGAGCTCAAAACCAACAGCAGACAGAAATTTGTACATCCAGAA 660
Qy 669 TACGAGCAGATTAACAGTCTCAACAAAAGAGTCAACAAACGATAGTCCCTTAATATCGGA 728
Db 661 TACGAGCAGATTAACAGTCTCAACAAAAGAGTCAACAAACGATAGTCCCTTAATATCGGA 720
Qy 729 TCTAGACCGTGGGTAGGGGTCAATCAGGAGGATAAGCATATACCTGGACCATTTGTAATAA 788
Db 721 TCTAGACCGTGGGTAGGGGTCAATCAGGAGGATAAGCATATACCTGGACCATTTGTAATAA 780
Qy 789 CTTGGAGATATCTTAATATGATAAACAAGTAATGGCAACTTATTTGACCGCGGGGATATTTT 848
Db 781 CTTGGAGATATCTTAATATGATAAACAAGTAATGGCAACTTATTTGACCGCGGGGATATTTT 840
Qy 849 AAATTCGAAACAGGGGAAAGCTCTGTAATGAGATCAGATGCACCCATAGACATTTGTGTG 908
Db 841 AAATTCGAAACAGGGGAAAGCTCTGTAATGAGATCAGATGCACCCATAGACATTTGTGTG 900

QY	909	TCTGAATGTTATACACCAATGAAGCATCCCAACGACAAACCATTTCAAAATGTGAAC	968
Db	901	TCTGAATGTTATACACCAATGAAGCATCCCAACGACAAACCATTTCAAAATGTGAAC	960
QY	969	AAAGTTACATATGGAATGCCCAAGTATATACAGGCAAAACACTTTAAAGCTGGCCACT	1028
Db	961	AAAGTTACATATGGAATGCCCAAGTATATACAGGCAAAACACTTTAAAGCTGGCCACT	1020
QY	1029	GGCATGAGGAATATACAGAAAGCAATCAGA	1061
Db	1021	GGCATGAGGAATGTACCAGAAAACCAATCAGA	1053
RESULT 7			
LOCUS	EIVY14060	1060 bp	RNA linear VRL 18-APR-2005
DEFINITION	Influenza A virus (A/equi 2/Bollinas/96 (H3N8)) gene for hemagglutinin HA1 subunit, isolate A/equi 2/Bollinas/96.		
ACCESSION	Y14060		
VERSION	Y14060.1	GI:2765373	
KEYWORDS	hemagglutinin HA1 subunit.		
SOURCE	Influenza A virus (A/equi 2/Bollinas/96 (H3N8))		
ORGANISM	Influenza A virus (A/equi 2/Bollinas/96 (H3N8))		
REFERENCE	1 (bases 1 to 1060)		
AUTHORS	Oxburgh, L., Akerblom, L., Fridberger, T., Klingeborn, B. and Linne, T.		
TITLE	Identification of two antigenically and genetically distinct lineages of H3N8 equine influenza virus in Sweden		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1060)		
AUTHORS	Oxburgh, L.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-JUN-1997) Oxburgh L.H., Veterinary Microbiology, Swedish University of Agricultural Sciences, Biomedical Centre Box 585, s-751 23 Uppsala, SWEDEN		
FEATURES	Location/Qualifiers		
source	1..1060	/organism="Influenza A virus (A/equi 2/Bollinas/96 (H3N8))"	
		/vireon	
		/mol_type="genomic RNA"	
		/serotype="H3N8"	
		/isolate="A/equi 2/Bollinas/96"	
		/db_xref="taxon:217819"	
mRNA	<22..1060		
CDS	22..1060	/codon_start=1	
		/product="hemagglutinin HA1 subunit"	
		/protein_id="CAA74388.1"	
		/db_xref="GI:2765373"	
		/db_xref="GOA:O56858"	
		/db_xref="InterPro:IPR000149"	
		/db_xref="InterPro:IPR001364"	
		/db_xref="UniProt/TrEMBL:O56858"	
		/translation="MKTILILLTHWVSONPTSGNTATILCLGHVANGTLVTKTI	
		TNDQIVTNATELVQSIGIKCINNSRVLDGNCTLIIDAMLDGPHCDVQFVNDLFP	
		IKRSFNCIPDIDYASLRISIVASSGTFEFTAEFTGWVTVQNGRSGACKRGSD	
		SFSLRNLTGNSYPTLVNTPMKNFDKLYIWGHPSSNQOQTELYIQESGRVT	
		VTKRSQQTIIPIIGSRPWRGSGRSIYWTIVKPGDILMINSGNLVAPRGYFLK	
		TGKSSVMRSDAPIDICASECITPNGSIPINDKPPQNVKVTYKPKVYIRQNTLKLATG	
		MNVPEKQIRGI"	
exon	22..1060		
		/number=1	
sig_peptide	22..67		
ORIGIN			
Query Match	98.2%	Score 1041.8;	DB 13; Length 1060;
Best Local Similarity	99.3%	Pred. No. 4.2e-231;	
Matches 1046; Conservative	0;	Mismatches 7;	Indels 0; Gaps 0;
QY	9	CAGGGGATATTTCTGTCAATCATGACAAACCATTTATTTGATCTACTGACCCATTGG	68

LOCUS	FLAHAH3D	1762 bp ss-RNA	linear	VRL 01-NOV-2004
DEFINITION	Influenza A virus (A/equine/Florida/1/93 (H3N8)) hemagglutinin precursor (HA) gene, complete cds.			
ACCESSION	L39916			
VERSION	L39916.1	GI:722407		
KEYWORDS				
SOURCE	Influenza A virus (A/equine/Florida/1/93 (H3N8))			
ORGANISM	Influenza A virus (A/equine/Florida/1/93 (H3N8))			
	Viruses; ssRNA negative-strand viruses; Orthomyxoviridae; Influenzavirus A.			
REFERENCE	1 (bases 1 to 1762)			
AUTHORS	Daly,J.M., Lai,A.C., Binns,M.M., Chambers,T.M., Barrandeguy,M. and Mumford,J.A.			
TITLE	Antigenic and genetic evolution of equine H3N8 influenza A viruses			
JOURNAL	J. Gen. Virol. 77 (Pt 4), 661-671 (1996)			
PUBLISHED	8627254			
FEATURES	Location/Qualifiers			
source	1. .1762			
	/organism="Influenza A virus			
	(A/equine/Florida/1/93 (H3N8))"			
	/mol_type="genomic RNA"			
	/strain="A/equine/Florida/1/93"			
	/serotype="H3N8"			
	/db_xref="taxon:39124"			
	/clone="2763"			
	/note="Equine-2"			
gene	30. .1727			
	/gene="HA"			
CDS	30. .1727			
	/gene="HA"			
	/codon_start=1			
	/product="hemagglutinin precursor"			
	/protein_id="AAB36978.1"			
	/db_xref="GI:722408"			
	/translation="MKTIILILLTHWVYQNPTSGNNTATLCIGHHAVANGTLVTKTI			
	TDQIEVATETALVQSISIGIKI CNNSYRVLDRNCTLIDAMLGDPHCDVFPYENWDLF			
	IERSSAFSNCPYDIDPYASLRSIVASSGTEFTAGFTWGTVQNGRSGACKRGASD			
	SFPSRLNMLTKSGNSYPTLANTWPNKNFKDLXIWGI:HPHSSNQEQTKLVIQBSGRVT			
	PLNKGSRQOITII PNIGSRPVRWGSGRISIVYTIKPGDILMINSGLNGLVAPRGYFKL			
	TGKSSVMSRSDPIDICVSECIITNGSISNKPQNVKNVYTGKCPKPIRNTKLKLAG			
	MRNVPEKIRGIFGAIAGFLENGWEVMDVGFGRYQNSGEGTQGAADLKSTQAAIDQI			
	NGKLNRIERTNPHQFIEKFESEVEGRIQDLEKYVEDTDIDLWSYNAELLVALENGH			
	TDILDAEMNKLFEKTRQLRENAEDMGGCFKIYHKCDNACIGISIRNGTYDHYIYRD			
	EALNNRFOIKSGVELKSGYKDWILWISFTISCFLLICVLLGFIWACQKGNIRNCIQT			
sig_peptide	30. .74			
	/gene="HA"			
	/note="putative"			
mat_peptide	75. .1061			
	/gene="HA"			
	/product="HA1"			
mat_peptide	1065. 1724			
	/gene="HA"			
	/product="HA2"			
ORIGIN				
Query Match	98.2%;	Score 1041.8;	DB 13;	Length 1762;
Best Local Similarity	98.9%;	Pred. No. 3.7e-231;		
Matches 1049;	Conservative 0;	Mismatches 12;	Indels 0;	Gaps 0;
Qy	1	AGCAAAAGCAGGGGATATTTCTGTCAATCATGAAGACAACCATATTTTGTACTACTGA	60	
Db	1	AGCAAAAGCAGGGGATATTTCTGTCAATCATGAAGACAACCATATTTTGTACTACTGA	60	
Qy	61	CCCATTTGGGTCTACAGTCAAAACCCACAGCTGGAACAACAACACAGCACCATTTGCTCGG	120	
Db	61	CCCATTTGGGTCTACAGTCAAAACCCACAGCTGGAACAACAACAGCACCATTTGCTCGG	120	
Qy	121	GACACCATGCAGTAGCAAAATGGACATTTGGTAAAAACAATAACTGATGACCAAAATGAGG	180	
Db	121	GACACCATGCAGTAGCAAAATGGACATTTGGTAAAAACAATAACTGATGACCAAAATGAGG	180	
Qy	181	TGCAAAATGCTACTGAATTAGTTTCAGAGCATTTTCAATAGGGAAAAATATGCAACACTCAT	240	

TITLE Identification of two antigenically and genetically distinct lineages of H3N8 equine influenza virus in Sweden
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1100)
 AUTHORS Oxburgh, L.H.
 TITLE Direct Submission
 JOURNAL Submitted (25-JUN-1997) Oxburgh L.H., Veterinary Microbiology, Swedish University of Agricultural Sciences, Biomedical Centre Box 585, s-751 23 Uppsala, SWEDEN
 FEATURES Location/Qualifiers
 source 1..1100
 /organism="Influenza A virus (A/equi 2/Soderala/94 (H3N8))"
 /viralion
 /mol_type="genomic RNA"
 /serotype="H3N8"
 /isolate="A/equi 2/Soderala/94"
 /db_xref="taxon:217821"
 <30..>1100
 30..>1100
 /codon_start=1
 /product="hemagglutinin HA1 subunit"
 /protein_id="CAA74386.1"
 /db_xref="GI:2765370"
 /db_xref="GOA:O56856"
 /db_xref="InterPro:IPR000149"
 /db_xref="InterPro:IPR001364"
 /db_xref="UniProt/TREMBL:O56856"
 translation="MKTTILILIHCVSQNPISGNNTATLCIGHAVANGTLVKTITDDIEVTNATELVQSISGKICNNISRVLDGRNCTLIDAMGDPHCDFOYENWDULIERSSAFNCPYDIPYASLSIVASLGLEFPAEGFTWTGVTQNSRSGACKRGSADVSTFRLNKLKNSGYPTLNWNNKFNPKLYIWIHPSNQOQTELYIOBSGRVTVTKRQOQTVNIGSPRPWVGSGRISYIWTIVKPGDILMINSNGLNVAIPRGVFLKLTGKSSVNRSDAPIDICVSECTINGSPNDKPFQNVKVTYKPKYIRQNTLKLATGMENVPEKQIRGIFGAIGFIENG"
 exon 30..1100
 /number=1
 sig_peptide 30..75
 ORIGIN
 Query Match 97.9%; Score 1039.2; DB 13; Length 1100;
 Best Local Similarity 98.8%; Pred. No. 1.7e-230;
 Matches 104; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 QY 2 GCAAAAGCAGGGGATATTTCTCAATCATGAAGAACCAACCATTTTGTATCTACTGAC 61
 DB 2 GCAAAAGCAGGGGATATTTCTCAATCATGAAGAACCAACCATTTTGTATCTACTGAT 61
 QY 62 CCATTGGGTCTACAGTCAAAACCCCAACAGTGGGAAACAACAGCCACATTTATGCTGGG 121
 DB 62 CCATTGGGTCTACAGTCAAAACCCCAACAGTGGGAAACAACAGCCACATTTATGCTGGG 121
 QY 122 ACACCATGCAGTAGCAATGGACATTTGGTAAACCAATTAAGTACCAATTTAGGT 181
 DB 122 ACACCATGCAGTAGCAATGGACATTTGGTAAACCAATTAAGTACCAATTTAGGT 181
 QY 182 GACAAATGCTACTGAAATAGTTTACAGCATTTCAATAGGGAATAATGCAACACTCATA 241
 DB 182 GACAAATGCTACTGAAATAGTTTACAGCATTTCAATAGGGAATAATGCAACACTCATA 241
 QY 242 TAAAGTTCTAGATGGAAGAATTCGACATTAATAGATGCAATGCTAGGAGACCCCACTG 301
 DB 242 TAGAGTTCTAGATGGAAGAATTCGACATTAATAGATGCAATGCTAGGAGACCCCACTG 301
 QY 302 TGATGCTTCCAGTATGAGATTTGGGACCTCTTCATAGAAAGAGCAGCGCTTCAGCAA 361
 DB 302 TGATGCTTCCAGTATGAGATTTGGGACCTCTTCATAGAAAGAGCAGCGCTTCAGCAA 361
 QY 362 TTGCTACCCATATGACATCCCTGACTATGATCGCTCGGTCATTGTAGCATCCTCAGG 421
 DB 362 TTGCTACCCATATGACATCCCTGACTATGATCGCTCGGTCATTGTAGCATCCTTAGG 421
 QY 422 AACATTAGAATTTCACAGCAGGGGATTCATATGGACAGGTGTCACTCAAAACCGGAAGAAG 481

Db 422 AACATTGGAATTCACAGCAGAGGGGATTCATCATGACAGGTGTCACTCAAAACCGGAAGAAG 481
 QY 482 TGGAGCCTGCAGAAAGGGGATCAGCCGATAGTTCTTTTAGCCGACTGAATTCGCTTAACAAA 541
 Db 482 TGGAGCCTGCAGAAAGGGGATCAGCCGATAGTTCTTTTAGCCGACTGAATTCGCTTAACAAA 541
 QY 542 ATCTGGAACCTCTTACCCCACTTGAATGTGACAATGCTTAACAAATTTTCGACAA 601
 Db 542 ATCTGGAACCTCTTACCCCACTTGAATGTGACAATGCTTAACAAATTTTCGACAA 601
 QY 602 ACTATACATCTGGGGGATTCATCACCCGAGCTCAACCAACAGCAGACAGATTTGTACAT 661
 Db 602 ACTATACATCTGGGGGATTCATCACCCGAGCTCAACCAACAGCAGACAGATTTGTACAT 661
 QY 662 CCAAGAATCAGGACGAGTAAACAGTCTCAACAAAAGAGTCAACAAAAGAGTCAACAAAAGTAAATCCCTTAA 721
 Db 662 CCAAGAATCAGGACGAGTAAACAGTCTCAACAAAAGAGTCAACAAAAGAGTCAACAAAAGTAAATCCCTTAA 721
 QY 722 TATCGGATCTAGACCGTGGGTTAGGGGTCAATCAGGCGAGGATAAGCATATATCTGACCAT 781
 Db 722 TATCGGATCTAGACCGTGGGTCAATCAGGCGAGGATAAGCATATATCTGACCAT 781
 QY 782 TGTAAACCTGGAGATATCTTAATGATAAACAAGTAATGGCAACTTAGTTGCAACCGCGGG 841
 Db 782 TGTAAACCTGGAGATATCTTAATGATAAACAAGTAATGGCAACTTAGTTGCAACCGCGGG 841
 QY 842 ATATTTTAAATTTGAAACACAGGAAAGCTCTGTAATGAGATCAGATGACCATCATAGACAT 901
 Db 842 ATATTTTAAATTTGAAACACAGGAAAGCTCTGTAATGAGATCAGATGACCATCATAGACAT 901
 QY 902 TTGTGTCTCTGAATGATTATACCAAAATGGAAGCATCCCCAACGACAAACCATTTTCAAAA 961
 Db 902 TTGTGTCTCTGAATGATTATACCAAAATGGAAGCATCCCCAACGACAAACCATTTTCAAAA 961
 QY 962 TGTGAACAAAGTTACATATGGAATGCCCAAGTATATATCAGGCAAAACACATTTAAAGCT 1021
 Db 962 TGTGAACAAAGTTACATATGGAATGCCCAAGTATATATCAGGCAAAACACATTTAAAGCT 1021
 QY 1022 GGCACCTGGGATGAGGAATATACCAAGAAAGCAATCAGA 1061
 Db 1022 GGCACCTGGGATGAGGAATATACCAAGAAAGCAATCAGA 1061
 FLAHAH3F 1762 bp ss-RNA linear VRL 01-NOV-2004
 LOCUS Influenza A virus (A/equine/Kentucky/1/91 (H3N8)) hemagglutinin
 DEFINITION precursor (HA) gene, complete cds.
 ACCESSION L39918
 VERSION L39918.1 GI:722411
 KEYWORDS Influenza A virus (A/equine/Kentucky/1/91 (H3N8))
 SOURCE Influenza A virus (A/equine/Kentucky/1/91 (H3N8))
 ORGANISM Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
 REFERENCE 1 (bases 1 to 1762)
 AUTHORS Daly, J.M., Lai, A.C., Binns, M.M., Chambers, T.M., Barrandeguy, M. and Mumford, J.A.
 TITLE Antigenic and genetic evolution of equine H3N8 influenza A viruses
 JOURNAL J. Gen. Virol. 77 (Pt 4), 661-671 (1996)
 PUBMED 8627254
 FEATURES Location/Qualifiers
 source 1..1762
 /organism="Influenza A virus (A/equine/Kentucky/1/91 (H3N8))"
 /mol_type="genomic RNA"
 /strain="A/equine/Kentucky/1/91"
 /serotype="H3N8"
 /db_xref="taxon:38828"
 /clone="2554"
 /note="Equine-2"
 30..1727
 /gene="HA"

CDS	30..1727	/gene="HA"	/codon_start=1	/product="hemagglutinin precursor"	/protein_id="AB36980.1"	/db_xref="GI:722412"	/translation="MKTTIILILLTHWYSONPTSGNNTATLCLGHHAIVANGTLVKTI TDDQTEVNATELVQSIIGIKCNNSYRLDGRNCTLIDAMLGDPHCDVQYENWDLF IERSFAFSKCPYDIPDYASLRSIVASSGTLFEAGFTWTGTQNGRSCKRGSAD SFPSLSNMLTKSGNSYPTLVNTPNNKNFDKLYIWTGTHPSSNEEOTKLYIQESGRVT VSTKSSQOTIIPNIGSRPMWQSGRISILYTIKPGDILMINSNGLVAPRGYFKLK TKSSWSDAPIDFIVDSBECTIPGNDPFPONVKNYTIKPKTIKQNTLUKLAG MKNVPEKIOIFGIAIGAFIENGWGMVDFQYQNSGEGTQQAOLDSTQAIDOI NGLKNRVITERNKEFQIEKSEVEGRIODLKEYIVDETKIDLWSYNAELLVLENQD TIDLDIAEMKNLFEKTRQLRENAEDMGCGCFKIYHKDCDNACIGSRNNGYDIHVIRD EALNNRFOIKGVELKSGYKDWILWISFAISFLICVLLGFIIMWACQKNGIRNICI"
	sig_peptide	/gene="HA"	/note="putative"				
	mat_peptide	75..1061	/gene="HA"	/product="HA1"			
	mat_peptide	1065..1724	/gene="HA"	/product="HA2"			
ORIGIN							
	Query Match	97.9%;	Score 1038.6;	DB 13;	Length 1762;		
	Best Local Similarity	98.7%;	Pred. No. 2.1e-230;	Mismatches 14;	Indels 0;	Gaps 0;	
	Matches 1047;	Conservative 0;					
QY	1	AGCAAAAGCAGGGGATATTCTGTCAATCATGACGAACAACCATATTATTGTATCTACTGA	60				
DB	1	AGCAAAAGCAGGGGATATTCTGTCAATCATGACGAACAACCATATTATTGTATCTACTGA	60				
QY	61	CCCATTTGGTCTACAGTCAAAACCCCAACCCAGTGGGAAACAACACGACCATATTATGTCGG	120				
DB	61	CCCATTTGGTCTACAGTCAAAACCCCAACCCAGTGGGAAACAACACGACCATATTATGTCGG	120				
QY	121	GACACATCAGTAGCAAAATGCAATGGTGTAAACAAACAATCACTGATGACCAAAATTCAGG	180				
DB	121	GACACATCAGTAGCAAAATGCAATGGTGTAAACAAACAATCACTGATGACCAAAATTCAGG	180				
QY	181	TGCAAAATGCTACTGAATTAGTTTCAGAGCATTTCAATAGGAGAAATATGCAACAACATCAT	240				
DB	181	TGCAAAATGCTACTGAATTAGTTTCAGAGCATTTCAATAGGAGAAATATGCAACAACATCAT	240				
QY	241	ATAAAGTTCTAGATGAAGAATAATGCACATTAATAGATGCAATGCTAGGAGACCCCACT	300				
DB	241	ATAAAGTTCTAGATGAAGAATAATGCACATTAATAGATGCAATGCTAGGAGACCCCACT	300				
QY	301	GTGATGCTCTCCAGTATGAAATTTGGGACCTCTTCATAGAAGAGCAGCGCTTCAGCA	360				
DB	301	GTGATGCTCTTCAGTATGAAATTTGGGACCTCTTCATAGAAGAGCAGCGCTTCAGCA	360				
QY	361	ATTGCTACCATATGACATCCCTGACTATGCAATGCAATGCTCCGCTCATTGTAGCATCTTCAG	420				
DB	361	ATTGCTACCATATGACATCCCTGACTATGCAATGCAATGCTCCGCTCATTGTAGCATCTTCAG	420				
QY	421	GAACTATTAGAAATTCACGACAGAGGATTCACATGGACAGGTGTCTACTCAAAACGGAAAGAA	480				
DB	421	GAACTATTAGAAATTCACGACAGAGGATTCACATGGACAGGTGTCTACTCAAAACGGAAAGAA	480				
QY	481	GTGGAGCTCGAAAAGGGGATCAGCGGATAGTTCCTTTTAGCCGACTGAATTTGGCTAACAA	540				
DB	481	GTGGAGCTCGAAAAGGGGATCAGCGGATAGTTCCTTTTAGCCGACTGAATTTGGCTAACAA	540				
QY	541	AATCTGGAACCTCTTACCCCAATTTGAATGTGACAAATGCCTAACAAATAAAATTTTCACA	600				
DB	541	AATCTGGAACCTCTTACCCCAATTTGAATGTGACAAATGCCTAACAAATAAAATTTTCACA	600				
QY	601	AACTATACATCTGGGGATTTCATCCCGAGCTCAAAACCAACAGCAGACAGAAATTTGTACA	660				

TDDIEVNTATVLOSISIGIKCNNSRVLDGRNCTLDIMLGDHPDCHVQYENWDLF
IERSATSNVCYDIPDYASLRISIVASSGLFETAGFTWTGVTQNGRSGACKRGSD
SFFRLNATKSGNSYPTLVNTPNNFNFKLYIWIHPPSSNOETKLVIOESGRVT
VTKRSQOTIIPNIGSRPWVQSGRISIVYIWKPDILMINSNGLVAPRGYFKLK
TGKSSVMSDAPIDICVSECITPNSISNDKPFQNVKNVYIGKPKYIRONTLKIATG
MRNVPEKQIR
30..74
/gene="HAI"
75..>1061
/gene="HAI"
/product="hemagglutinin"

sig_peptide
mat_peptide

ORIGIN

Query Match 97.6%; Score 1035.4; DB 13; Length 1061;
Best Local Similarity 98.5%; Pred. No. 1.3e-229;
Matches 1045; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 1 AGCAAAAGCAGGGGATATTTCTGTCATCATGAAGACACCAATATTTGATCTACTGA 60
Db 1 AGCAAAAGCAGGGGATATTTCTGTCATCATGAAGACACCAATATTTGATCTACTGA 60

Qy 61 CCCATTGGGCTACAGTCAAAACCAACCAAGTGAAGACACACAGCCACATTTGCTCGG 120
Db 61 CCCATTGGGCTACAGTCAAAACCAACCAAGTGAAGACACACAGCCACATTTGCTCGG 120

Qy 121 GACACCATGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 180
Db 121 GACACCATGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 180

Qy 181 TGACAAATGCTACTGAATTTAGTTCAGAGCATTTCAATAGGAAATATGCAACAACTCAT 240
Db 181 TGACAAATGCTACTGAATTTAGTTCAGAGCATTTCAATAGGAAATATGCAACAACTCAT 240

Qy 241 ATAAAGTCTAGATGGAGAAATGGACATTAATAGATGCAATGCTAGGAGACCCCACT 300
Db 241 ATAGAGTCTAGATGGAGAAATGGACATTAATAGATGCAATGCTAGGAGACCCCACT 300

Qy 301 GTGATGCTTCCAGTATGAGAAATGGGACCTCTTCATAGAAAGACAGCGCTTTTCAGCA 360
Db 301 GTGATGCTTCCAGTATGAGAAATGGGACCTCTTCATAGAAAGACAGCGCTTTTCAGCA 360

Qy 361 ATTGCTACCATATGACATCCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 420
Db 361 ATTGCTACCATATGACATCCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 420

Qy 421 GACATTAGAAATCACAGCAGAGGATTCATGACAGAGTGTGCTCAAAACGAGAA 480
Db 421 GACATTAGAAATCACAGCAGAGGATTCATGACAGAGTGTGCTCAAAACGAGAA 480

Qy 481 GTGAGGCTGCAAAAGGGGATCAGCCGATAGTTTCTTTAGCCGACTGAAATGGCTAACAA 540
Db 481 GTGAGGCTGCAAAAGGGGATCAGCCGATAGTTTCTTTAGCCGACTGAAATGGCTAACAA 540

Qy 541 AATCTGAAATCTTACCCACATTTGAATGTGAATGTGAATGTGAATGTGAATGTGAATGTGA 600
Db 541 AATCTGAAATCTTACCCACATTTGAATGTGAATGTGAATGTGAATGTGAATGTGAATGTGA 600

Qy 601 AACTATACATCTGGGGATTCATCACCAGGCTCAAAACCAAGAGCAGACAGAAATTTGACA 660
Db 601 AACTATACATCTGGGGATTCATCACCAGGCTCAAAACCAAGAGCAGACAGAAATTTGACA 660

Qy 661 TCAAGAAATCAGGACGAGTAACTGCTCAACAAAGAAAGTCAACAAACGATAGTCCCTA 720
Db 661 TCAAGAAATCAGGACGAGTAACTGCTCAACAAAGAAAGTCAACAAACGATAGTCCCTA 720

Qy 721 ATATCGGATCTAGACCGTGGGTTAGGGGTCAATCAGGAGGATTAAGCATATCTGGACCA 780
Db 721 ATATCGGATCTAGACCGTGGGTTAGGGGTCAATCAGGAGGATTAAGCATATCTGGACCA 780

Qy 781 TTGTAAACCTGGAGATATCTTAATGATAACAGTAAATGGCACTTGTGTCACCGCGG 840
Db 781 TTGTAAACCTGGAGATATCTTAATGATAACAGTAAATGGCACTTGTGTCACCGCGG 840

Qy 841 GATATTTTAAATTTGAAACACAGGAAAAAGCTCTGTAATGAGATCAGATGCACCCATAGACA 900
Db 841 GATATTTTAAATTTGAAACACAGGAAAAAGCTCTGTAATGAGATCAGATGCACCCATAGACA 900

Qy 901 TTTGTGTCTGTAATGATTAACACCAATGAAGCATCCCAACGACAAACCACTTTTAAAGC 960
Db 901 TTTGTGTCTGTAATGATTAACACCAATGAAGCATCTCCACGACAAACCACTTTTAAAGC 960

Qy 961 ATGTGAACAAAGTTTACATATGGAATGCCCCCAAGTATATCAGGCAAAACCACTTTAAAGC 1020
Db 961 ATGTGAACAAAGTTTACATATGGAATGCCCCCAAGTATATCAGGCAAAACCACTTTAAAGC 1020

Qy 1021 TGGCCACTGGGATGAGGAATATATACGAGAAAGCAAAATCAGA 1061
Db 1021 TGGCCACTGGGATGAGGAATATATACGAGAAAGCAAAATCAGA 1061

RESULT 12
BD244631
LOCUS 1762 bp DNA linear PAT 17-JUL-2003
DEFINITION Low temperature-adaptable equine influenza virus.
ACCESSION BD244631
VERSION BD244631.1 GI:33054401
KEYWORDS JP 2002522078-A/7.
SOURCE Equine influenza virus H3N8
ORGANISM Equine influenza virus H3N8
Viruses; serNA negative-strand viruses; Orthomyxoviridae;
Influenzavirus A.
REFERENCE 1 (bases 1 to 1762)
AUTHORS Dowling,P.W. and Youngner,J.S.
TITLE Low temperature-adaptable equine influenza virus
JOURNAL Patent: JP 2002522078-A 7 23-JUL-2002;
THE UNIVERSITY OF PITTSBURGH OF THE COMMONWEALTH SYSTEM OF HIGHER
EDUCATION
COMMENT OS Equine influenza virus H3N8
PN JP 2002522078-A/7
PD 23-JUL-2002
PF 12-AUG-1999 JP 2000565137
PR 13-AUG-1998 US 09/133921
PI PATRICIA W DOWLING,JULIUS S YOUNGNER
PC C12N15/09,A61K39/145,A61P31/16,C07K14/11,C12N7/04//(C12N7/04,
C12R1:92), PC
PC C12N15/00
CC Low temperature-adaptable equine influenza virus FH Key
FT CDS Location/Qualifiers
source 1..1762
Location/Qualifiers
1..1762
/organism="Equine influenza virus H3N8"
/mol_type="genomic DNA"
/db_xref="taxon:31660"

FEATURES
ORIGIN

Query Match 97.4%; Score 1033.8; DB 6; Length 1762;
Best Local Similarity 98.4%; Pred. No. 2.7e-229;
Matches 1044; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 1 AGCAAAAGCAGGGGATATTTCTGTCATCATGAAGACCAACCAATATTTGATCTACTGA 60
Db 1 AGCAAAAGCAGGGGATATTTCTGTCATCATGAAGACCAACCAATATTTGATCTACTGA 60

Qy 61 CCCATTGGGCTACAGTCAAAACCAACCAAGTGAAGACACAGCCACATTTGCTCGG 120
Db 61 CCCATTGGGCTACAGTCAAAACCAACCAAGTGAAGACCAACCAATATTTGCTCGG 120

Qy 121 GACACCATGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 180
Db 121 GACACCATGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 180

Qy 181 TGACAAATGCTACTGAATTTAGTTCAGAGCATTTCAATAGGAAATATGCAACAACTCAT 240
Db 181 TGACAAATGCTACTGAATTTAGTTCAGAGCATTTCAATAGGAAATATGCAACAACTCAT 240


```
QY 961 ATGTGAACAAAGTTACATATGGAATGCCCCAAGTATATCAGGCAAAACACTTTAAAGC 1020
Db 961 ATGTGAACAAAGTTACATATGGAATGCCCCAAGTATATCAGGCAAAACACTTTAAAGC 1020
QY 1021 TGGCCACTGGGATGAGGAAATACAGAAAGCAAAATCAGA 1061
Db 1021 TGGCCACTGGGATGAGGAAATACAGAAAGCAAAATCAGA 1061

RESULT 14
AR343239 1762 bp DNA linear PAT 17-AUG-2003
LOCUS AR343239
DEFINITION Sequence 10 from patent US 6579528.
ACCESSION AR343239
VERSION AR343239.1 GI:33738757
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
AUTHORS Dowling, P.W. and Youngner, J.S.
TITLE Cold-adapted equine influenza viruses
JOURNAL Patent: US 6579528-A 10 17-JUN-2003;
The University of Pittsburgh - of the Commonwealth System of Higher
Education; Pittsburgh, PA
FEATURES
source Location/Qualifiers
1..1762
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 97.4%; Score 1033.8; DB 6; Length 1762;
Best Local Similarity 98.4%; Pred. No. 2.7e-229;
Matches 1044; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 AGCAAAAGCAGGGGATATTTCTGTCATCATGAGACAAACCATTTATTTGATCTACTGA 60
Db 1 AGCAAAAGCAGGGGATATTTCTGTCATCATGAGACAAACCATTTATTTGATCTACTGA 60
QY 61 CCCATTGGGTCTACAGTCAAAACCCCAACAGTGGAAACAACACAGCCACATTTATGTCGG 120
Db 61 CCCATTGGGTCTACAGTCAAAACCCCAACAGTGGAAACAACACAGCCACATTTATGTCGG 120
QY 121 GACACCATGAGTACAGTAAATGGAACATTTGGTAAACCAATTAATGATGACCAATTTGAGG 180
Db 121 GACACCATGAGTACAGTAAATGGAACATTTGGTAAACCAATTAATGATGACCAATTTGAGG 180
QY 181 TGACAAATGCTACTGAAATTAGTTCAGAGCATTTTCAATAGGGAATAATGCAACAACTCAT 240
Db 181 TGACAAATGCTACTGAAATTAGTTCAGAGCATTTTCAATAGGGAATAATGCAACAACTCAT 240
QY 241 ATAAAGTTCATAGTGAAGAAATTTGCAATTATAGATGCAATGCTAGGAGACCCCACT 300
Db 241 ATAGAGTTCATAGTGAAGAAATTTGCAATTATAGATGCAATGCTAGGAGACCCCACT 300
QY 301 GTGATGCTTCCAGTATGAGAAATTTGGACCTTTCATAGAAAGACAGCGCTTTCAGCA 360
Db 301 GTGATGCTTTCAGTATGAGAAATTTGGACCTTTCATAGAAAGACAGCGCTTTCAGCA 360
QY 361 ATTGCTACCCATATGACATCCCTGACATATGATCGCTCCGCTCCATTTAGCATCCTCAG 420
Db 361 GTTGCTACCCATATGACATCCCTGACATATGATCGCTCCGCTCCATTTAGCATCCTCAG 420
QY 421 GAACATTAGAAATTCACAGCAGAGGATTCATGGAACAGGTGTCTCTCAAAAACGGAAGAA 480
Db 421 GAACATTAGAAATTCACAGCAGAGGATTCATGGAACAGGTGTCTCTCAAAAACGGAAGAA 480
QY 481 GTGAGCCTGCAAAAGGGATTCAGCGATGTTTCTTTAGCCGACTGAATTTGCTTACAA 540
Db 481 GTGAGCCTGCAAAAGGGATTCAGCGATGTTTCTTTAGCCGACTGAATTTGCTTACAA 540
QY 541 AATCTGAAACTCTTACCCCAACATTTGAATGTGACAAATGTCATTAATAAATAATTTTCGACA 600

QY 541 AATCTGAAACTCTTACCCCAACATTTGAATGTGACAAATGTCATTAATAAATAATTTTCGACA 600
QY 601 AACTATATCATCTGGGGATTCATCACCAGAGCTCAAAACCAACAGCAGACAGAAATTTGTACA 660
Db 601 AACTATATCATCTGGGGATTCATCACCAGAGCTCAAAACCAACAGCAGACAGAAATTTGTACA 660
QY 661 TCCAAAGAAATCAGGACAGTAACTCTCAACAAAGAAAGTCAACAAACAGATAGTCCCTTA 720
Db 661 TCCAAAGAAATCAGGACAGTAACTCTCAACAAAGAAAGTCAACAAACAGATAGTCCCTTA 720
QY 721 ATATCGGATCTAGACCGTGGGTAGGGTCAATCAGGACAGGATAGCATATATCTGGACCA 780
Db 721 ACATCGGATCTAGACCGTGGGTAGGGTCAATCAGGACAGGATAGCATATATCTGGACCA 780
QY 781 TTGTAAACCTCGAGATATCTTAATGATAACAGTAATGGCACTTATGTCACCGCGGG 840
Db 781 TTGTAAACCTCGAGATATCTTAATGATAACAGTAATGGCACTTATGTCACCGCGGG 840
QY 841 GATATTTTAAATTTGAAACAGGGAAGTCTCTGTAATGAGATCAGATGCAACCCATAGACA 900
Db 841 GATATTTTAAATTTGAAACAGGGAAGTCTCTGTAATGAGATCAGATGCAACCCATAGACA 900
QY 901 TTTGTGTCTGTAATGTAATACCAATGGAAGCATCCCCAACAGCAAAACCATTTCAAA 960
Db 901 TTTGTGTCTGTAATGTAATACCAATGGAAGCATCCCCAACAGCAAAACCATTTCAAA 960
QY 961 ATGTGAACAAAGTTACATATGGAATAATGCCCCAAGTATATCAGGCAAAACACTTTAAAGC 1020
Db 961 ATGTGAACAAAGTTACATATGGAATAATGCCCCAAGTATATCAGGCAAAACACTTTAAAGC 1020
QY 1021 TGGCCACTGGGATGAGGAAATATACGAAAGCAAAATCAGA 1061
Db 1021 TGGCCACTGGGATGAGGAAATATACGAAAGCAAAATCAGA 1061

RESULT 15
AR455506 1762 bp DNA linear PAT 20-FEB-2004
LOCUS AR455506
DEFINITION Sequence 10 from patent US 6685946.
ACCESSION AR455506
VERSION AR455506.1 GI:42690326
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
AUTHORS Dowling, P.W. and Youngner, J.S.
TITLE Cold-adapted equine influenza viruses
JOURNAL Patent: US 6685946-A 10 03-FEB-2004;
The University of Pittsburgh-of the Commonwealth System of Higher
Education; Pittsburgh, PA
FEATURES
source Location/Qualifiers
1..1762
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 97.4%; Score 1033.8; DB 6; Length 1762;
Best Local Similarity 98.4%; Pred. No. 2.7e-229;
Matches 1044; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 AGCAAAAGCAGGGGATATTTCTGTCATCATGAGACAAACCATTTATTTGATCTACTGA 60
Db 1 AGCAAAAGCAGGGGATATTTCTGTCATCATGAGACAAACCATTTATTTGATCTACTGA 60
QY 61 CCCATTGGGTCTACAGTCAAAACCCCAACAGTGGAAACAACACAGCCACATTTATGTCGG 120
Db 61 CCCATTGGGTCTACAGTCAAAACCCCAACAGTGGAAACAACACAGCCACATTTATGTCGG 120
QY 121 GACACCATGAGTACAGTAAATGGAACATTTGGTAAACCAATTAATGATGACCAATTTGAGG 180
Db 121 GACACCATGAGTACAGTAAATGGAACATTTGGTAAACCAATTAATGATGACCAATTTGAGG 180
QY 181 TGACAAATGCTACTGAAATTAGTTCAGAGCATTTTCAATAGGGAATAATGCAACAACTCAT 240
Db 181 TGACAAATGCTACTGAAATTAGTTCAGAGCATTTTCAATAGGGAATAATGCAACAACTCAT 240
QY 241 ATAAAGTTCATAGTGAAGAAATTTGCAATTATAGATGCAATGCTAGGAGACCCCACT 300
Db 241 ATAGAGTTCATAGTGAAGAAATTTGCAATTATAGATGCAATGCTAGGAGACCCCACT 300
QY 301 GTGATGCTTCCAGTATGAGAAATTTGGACCTTTCATAGAAAGACAGCGCTTTCAGCA 360
Db 301 GTGATGCTTTCAGTATGAGAAATTTGGACCTTTCATAGAAAGACAGCGCTTTCAGCA 360
QY 361 ATTGCTACCCATATGACATCCCTGACATATGATCGCTCCGCTCCATTTAGCATCCTCAG 420
Db 361 GTTGCTACCCATATGACATCCCTGACATATGATCGCTCCGCTCCATTTAGCATCCTCAG 420
QY 421 GAACATTAGAAATTCACAGCAGAGGATTCATGGAACAGGTGTCTCTCAAAAACGGAAGAA 480
Db 421 GAACATTAGAAATTCACAGCAGAGGATTCATGGAACAGGTGTCTCTCAAAAACGGAAGAA 480
QY 481 GTGAGCCTGCAAAAGGGATTCAGCGATGTTTCTTTAGCCGACTGAATTTGCTTACAA 540
Db 481 GTGAGCCTGCAAAAGGGATTCAGCGATGTTTCTTTAGCCGACTGAATTTGCTTACAA 540
QY 541 AATCTGAAACTCTTACCCCAACATTTGAATGTGACAAATGTCATTAATAAATAATTTTCGACA 600
```

181	Qy	TGCAAAATGCTACTGAAATAGTTTCAGAGCATTTCAATAGGGAATAATGCAACCACTCAT	240
181	Db	TGCAAAATGCTACTGAAATAGTTTCAGAGCATTTCAATAGGGAATAATGCAACCACTCAT	240
241	Qy	ATAAAGTTCTAGATGGAAGAAATTCACATTAATAGATGCAATGCTAGGAGACCCCACT	300
241	Db	ATAAGTTCTAGATGGAAGAAATTCACATTAATAGATGCAATGCTAGGAGACCCCACT	300
301	Qy	GTGATGTTCTCAGTATAGAAATTTGGACCTCTTCATAGAAAGACGCGCTTTCAGCA	360
301	Db	GTGATGTTCTCAGTATAGAAATTTGGACCTCTTCATAGAAAGACGCGCTTTCAGCA	360
361	Qy	ATTGCTTACCCATATGACATCCCTGACTATGCAATCGCTCCGGTCCATTTGTAGCATCTCTAG	420
361	Db	GTTGCTTACCCATATGACATCCCTGACTATGCAATCGCTCCGGTCCATTTGTAGCATCTCTAG	420
421	Qy	GAACATTTAGAAATTCACAGCAGAGGGATTCACATGGGACAGGTGTCACTCAAAACGGGAAGAA	480
421	Db	GAACATTTGGAATTCACAGCAGAGGGATTCACATGGGACAGGTGTCACTCAAAACGGGAAGAA	480
481	Qy	GTGAGCCTCGCAAAAGGGGATCAGCGATAGTTTCTTTAGCCGACGTGAATTGCTTAACAA	540
481	Db	GTGGATCTCGCAAAAGGGAAATCAGCGATAGTTTCTTTAGCCGACGTGAATTTGCTTAACAG	540
541	Qy	AATCTGGAAACTCTTATACCCACATTTGAATGTGACAAATGCCTAAACAATAAAAAATTCGACA	600
541	Db	AATCTGGAAACTCTTATACCCACATTTGAATGTGACAAATGCCTAAACAATAAAAAATTCGACA	600
601	Qy	AACATATACATCTGGGGATTCATCACCAGAGCTCAAAACCAACAGCAGACGAAATTTGTACA	660
601	Db	AACATAACATCTGGGGATTCATCACCAGAGCTCAAAACCAAGAGCAGACAAAATTTGTACA	660
661	Qy	TCCAAGAATCAGGACGAGTAAACAGTCTCAACAAAAAGAAAGTCAACAAACGATAGTCCCTTA	720
661	Db	TCCAAGNATCAGACGNGTAACAGTCTCAACAAAAGAAAGTCAACAAACNATTAATCCCTTA	720
721	Qy	ATATCGGAATCAGACCGTGGGTTAGGGTCAATCAGGCAGGATTAAGCATATATCTGCAACCA	780
721	Db	ACATCGGAATCAGACCGTGGGTCAATCAGGCAGGATTAAGCATATATCTGCAACCA	780
781	Qy	TTGTAAACCTGGAGATATCCTTAATGATAACAGTAAATGGCAACTTGTAGTTCACACCGGG	840
781	Db	TTGTAAACCTGGAGATATCCTAACGATAACAGTAAATGGCAACTTGTAGTTCGACCCGGG	840
841	Qy	GATATTTTAAATTTGAAACAGGAAAAGCTCTCTGAATGAGATCAGATGCAACCCATAGACA	900
841	Db	GATATTTTAAATTTGAAACAGGAAAAGCTCTGTGAATGAGATCAGATGCAACCCATAGACA	900
901	Qy	TTTTGTGTCGTAATGTATTTACACCAATGGAAGCATCCCCAAGCAACCAATTTCAA	960
901	Db	TTTTGTGTCGTAATGTATTTACACCAATGGAAGCATCCCCAAGCAACCAATTTCAA	960
961	Qy	ATGTGAACAAAGTTTACATATGGAATAATGCCCAAGTATATCAGGCAAAAACACTTTAAAGC	1020
961	Db	ATGTGAACAAAGTTTACATATGGAATAATGCCCAAGTATATCAGGCAAAAACACTTTAAAGC	1020
1021	Qy	TGGCCACTGGGATGAGGAATATATCCAGAAAGCAAAATCAGA	1061
1021	Db	TGGCCACTGGGATGAGGAATGTATCCAGAAAGCAAAATCAGA	1061

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2005, 23:58:01 ; Search time 678 Seconds
(without alignments)
10429.556 Million cell updates/sec

Title: US-10-826-929A-1
Perfect score: 1061
Sequence: 1 agcaaaagcggggatattt.....taccagaaaagcaaatcaga 1061

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_21.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*
- 14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1061	100.0	1061	14	Adw87317 Equine-2
2	1033.8	97.4	1762	3	Aaz50976 Cold-adap
3	1032.2	97.3	1762	3	Aaz50975 Wild type
4	1019.2	96.1	1698	2	Aaz30213 Sequence
5	1018.4	96.0	1762	2	Aaz30212 Sequence
6	987.4	93.1	1788	1	Aan71067 Equine HA (A
7	987.4	93.1	1797	1	Aan71067 Sequence
8	984	92.7	1698	3	Aaz30211 Equine he
9	984	92.7	1698	3	Aaz47007 Equine in
10	976.2	92.0	1762	2	Aaz30211 Equine in
11	947.2	89.3	1698	2	Aav49391 EIV HA (A
12	607.2	57.2	1777	2	Aaz72855 Full. leng
13	601.6	56.7	1765	6	Aba93937 Influenza
14	598.4	56.4	1764	6	Aba93944 Influenza
15	585.6	55.2	1762	1	Aan70642 Sequence
16	579.2	54.6	1762	14	Ady27535 Hemagglut
17	574.4	54.1	1762	14	Ady27539 Hemagglut
18	569.6	53.7	1755	14	Ady27533 Hemagglut
19	568	53.5	1762	14	Ady27541 Hemagglut

20	566.4	53.4	1762	14	ADY27537	Ady27537 Hemagglut
21	566.2	53.4	1091	6	ABQ82724	Abq82724 Influenza
22	564.8	53.2	1745	14	ADY27531	Ady27531 Hemagglut
23	558.2	52.6	1653	6	ABQ82723	Abq82723 Influenza
24	556	52.4	987	6	ABQ82725	Abq82725 Influenza
25	554.4	52.3	987	6	ABQ82725	Abq82725 Influenza
26	549.2	51.8	1793	2	AAT59212	Aat59212 Influenza
27	549.2	51.8	1793	2	AAT59212	Aat59212 Influenza
28	549.2	51.8	1793	5	AAD09586	Aad09586 Influenza
29	546.4	51.5	988	12	ADOL5237	Adol5237 Influenza
30	544.4	51.3	1757	2	AAT59216	Aat59216 Influenza
31	544.4	51.3	1757	2	AAO00777	Aao00777 Influenza
32	544.4	51.3	1757	5	AAD09590	Aad09590 Influenza
33	541.2	51.0	1701	2	AAV49298	Aav49298 SIV strai
34	541.2	51.0	1757	2	AAT59219	Aat59219 Influenza
35	541.2	51.0	1757	2	AAO00780	Aao00780 Influenza
36	541.2	51.0	1757	5	AAD09593	Aad09593 Influenza
37	150.4	14.2	1809	1	AAN71066	Aan71066 Sequence
38	150.4	14.2	1809	2	AAQ04596	Aaq04596 Equine he
39	148.2	14.0	1759	2	AAQ29110	Aaq29110 EIV HA (A
40	133.4	12.6	2005	3	AAA75002	Aaa75002 Nucleotid
41	133.4	12.6	4610	3	AAA75005	Aaa75005 Nucleotid
42	133.4	12.6	4930	3	AAA75000	Aaa75000 Nucleotid
43	127.2	12.0	1721	2	ADH29821	Adh29821 Swinepox
44	127.2	12.0	1721	2	AAV26247	Aav26247 Genomic D
45	125.2	11.8	1711	6	AAL50113	Aal50113 Recombina

ALIGNMENTS

RESULT 1

ADW87317
ID ADW87317 standard; DNA; 1061 BP.

XX AC ADW87317;

XX DT 21-APR-2005 (first entry)

XX DE Equine-2 influenza virus hemagglutinin protein HA1 coding sequence.

XX KW immunostimulant; vaccine; immunization; hemagglutinin; gene; ds.

XX OS Influenza virus.

XX FH Key Location/Qualifiers
FT CDS 30..1061

FT FT /*tag= b

FT FT /product= "hemagglutinin"

FT FT /note= "no stop codon given"

FT FT /transl_except= (pos:369..377,aa:ProAsp)

FT FT /note= "a one codon insertion alters the reading frame"

FT FT sig_peptide 30..74

FT FT mat_peptide 75..1061

FT FT /*tag= c

XX US2005032732-A1.

XX 10-FEB-2005.

XX PF 16-APR-2004; 2004US-00826929.

XX PR 15-MAY-2003; 2003US-0470843P.

XX PA (LAI/A/) LAI A.

XX PI Lai A;

XX XX WPI; 2005-151686/16.

XX DR P-PSDB; ADW87316.

XX DR GENBANK; AF197241.

XX XX

CC encoding cold-adapted equine influenza virus H3N8 haemagglutinin (HA)
CC protein denoted as PeiwhA565. This sequence is a modified form of the
CC wild type sequence
XX
SQ Sequence 1762 BP; 641 A; 333 C; 381 G; 407 T; 0 U; 0 Other;
Query Match 97.4%; Score 1033.8; DB 3; Length 1762;
Best Local Similarity 98.4%; Pred. No. 1.4e-281;
Matches 1044; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 1 AGCAAAAGCAGGGATATTTCTGTCATCATGAAGACCAACCATTTATTTGTACTACTGA 60
DB 1 AGCAAAAGCAGGGATATTTCTGTCATCATGAAGACCAACCATTTATTTGTACTACTGA 60
QY 61 CCCATTGGGTCTACAGTCAAAACCCAAACCAACCAAGTGGCAACACACAGCCACATTTGTCTGG 120
DB 61 CCCATTGGGTCTACAGTCAAAACCCAAACCAACCAAGTGGCAACACACAGCCACATTTGTCTGG 120
QY 121 GACACCATGAGTAGCAAAATGGAAATTTGGTAAACCAATTAACCTGATGACCAAAATTGAGG 180
DB 121 GACACCATGAGTAGCAAAATGGAAATTTGGTAAACCAATTAACCTGATGACCAAAATTGAGG 180
QY 181 TGCAAAATGCTACTGAATTTAGTTCAGAGCATTTCAATAGGAAATATGCAAACTCAT 240
DB 181 TGCAAAATGCTACTGAATTTAGTTCAGAGCATTTCAATAGGAAATATGCAAACTCAT 240
QY 241 ATAAAGTTCTAGATGGAAGAAATTCACATTAATAGATGCAATCTAGGAGACCCCACT 300
DB 241 ATAGAGTTCTAGATGGAAGAAATTCACATTAATAGATGCAATCTAGGAGACCCCACT 300
QY 301 GTGATGTTCTTCCAGTATGAGAAATTTGGGACCTCTTCATAGAAAGCAGCGCTTTTCAGCA 360
DB 301 GTGATGTTCTTCCAGTATGAGAAATTTGGGACCTCTTCATAGAAAGCAGCGCTTTTCAGCA 360
QY 361 ATTGCTACCATATGACATCCCTGACATATGATCGCTCCGGTCCATTTGTAGCATCTCTAG 420
DB 361 GTTGTCTACCATATGACATCCCTGACATATGATCGCTCCGGTCCATTTGTAGCATCTCTAG 420
QY 421 GAACATTTAGAAATTCACAGCAGAGGATTCACATGACAGAGTGTCACCTCAAAACGGAGAA 480
DB 421 GAACATTTAGAAATTCACAGCAGAGGATTCACATGACAGAGTGTCACCTCAAAACGGAGAA 480
QY 481 GTGAGGCTCGCAAAAGGGATTCAGCGATAGTTTCTTTAGCCGACTGAATTTGGCTAACAA 540
DB 481 GTGAGGCTCGCAAAAGGGATTCAGCGATAGTTTCTTTAGCCGACTGAATTTGGCTAACAA 540
QY 541 AATCTGAAACTCTTACCCACATTTGAATGTGACAAATGCCTTAACAATAAATAATTTGACA 600
DB 541 AATCTGAAACTCTTACCCACATTTGAATGTGACAAATGCCTTAACAATAAATAATTTGACA 600
QY 601 AACTATACATCTGGGGATTCATCACCAGGCTCAACACCAACAGCAGCAGAGATTTGTACA 660
DB 601 AACTATACATCTGGGGATTCATCACCAGGCTCAACACCAACAGCAGCAGAGATTTGTACA 660
QY 661 TCCAAGATCAGGACGAGTAACTCTCAACAAAGAGAGTCAACAAACGATAGTCCCTTA 720
DB 661 TCCAAGATCAGGACGAGTAACTCTCAACAAAGAGAGTCAACAAACGATAGTCCCTTA 720
QY 721 ATATCGGATCTAGACCGTGGGTAGGGTCAATCAGCAGGATGAAGATATCTGGACCA 780
DB 721 ACATCGGATCTAGACCGTGGGTAGGGTCAATCAGCAGGATGAAGATATCTGGACCA 780
QY 781 TTGTAAACCTGGAGATATCTTAATGATAACAGTAAATGCACTTGTGACCCGGGG 840
DB 781 TTGTAAACCTGGAGATATCTTAATGATAACAGTAAATGCACTTGTGACCCGGGG 840
QY 841 GATATTTAAATTCAAAACAGGGAAGAGTCTGTAAATGAGATCAGATGACACCATAGACA 900
DB 841 GATATTTAAATTCAAAACAGGGAAGAGTCTGTAAATGAGATCAGATGACACCATAGACA 900
QY 901 TTGTGTGTCTGAATGATTTACCAATGGAAGCATCTCCCAACGACAAACCATTTCAA 960
DB 901 TTGTGTGTCTGAATGATTTACCAATGGAAGCATCTCCCAACGACAAACCATTTCAA 960

QY 961 ATGTGAACAAGTTTACATATGGAATGCCCCCAAGTATATCAGGCAAAACACTTTAAAGC 1020
DB 961 ATGTGAACAAGTTTACATATGGAATGCCCCCAAGTATATCAGGCAAAACACTTTAAAGC 1020
QY 1021 TGGCCACTGGGATGAGGAATATACAGAAAGCAAAATCAGA 1061
DB 1021 TGGCCACTGGGATGAGGAATGTACCAGAAAGCAAAATCAGA 1061
RESULT 3
AAZ50975
ID AAZ50975 standard; DNA; 1762 BP.
XX
AC AAZ50975;
XX
DT 05-JUN-2000 (first entry)
XX
DE Wild type equine influenza virus H3N8 haemagglutinin protein DNA.
XX
KW Haemagglutinin protein; wild type HA protein; horse; cold-adaptation;
KW reassortant virus; temperature sensitivity; dominant interference;
KW attenuation; antiviral; vaccine; prevention; treatment;
KW influenza A virus infection; ds.
XX
OS Equine influenza virus H3N8.
XX
FH Key Location/Qualifiers
FT CDS 30..1727
FT /*tag= a
FT /product= "HA protein"
FT /note= "The coding region without the stop codon is
FT specifically claimed"
XX
PN WO200009702-A1.
XX
PD 24-FEB-2000.
XX
PF 12-AUG-1999; 99WO-US018583.
XX
PR 13-AUG-1998; 98US-00133921.
XX
PA (UVP1-) UNIV PITTSBURGH.
XX
PI Dowling PW, Youngner JS;
XX
DR WPI; 2000-224339/19.
XX P-PSDB; AAY70056.
XX
PT New cold-adapted equine influenza viruses and reassortant viruses used as
PT vaccines for treating influenza infections in animals, particularly
PT horses, have a phenotype such as temperature sensitivity or dominant
PT interference.
XX
PS Claim 8; Page 82-85; 127pp; English.
XX
CC The patent discloses experimentally generated cold-adapted equine
CC influenza viruses, and reassortant influenza A viruses comprising at
CC least one genome segment of the cold-adapted virus, which confers at
CC least one identifying phenotype selected from cold-adaptation,
CC temperature sensitivity, dominant interference and attenuation. These
CC viruses are used in therapeutic compositions e.g. vaccines for preventing
CC or treating infections caused by influenza A viruses in animals,
CC particularly horses. The present sequence is a DNA (neiwthA1762) encoding
CC wild type equine influenza virus H3N8 haemagglutinin (HA) protein denoted
CC as PeiwhA565. This sequence is modified to generate cold-adapted equine
CC influenza virus
XX
SQ Sequence 1762 BP; 639 A; 334 C; 383 G; 406 T; 0 U; 0 Other;
Query Match 97.3%; Score 1032.2; DB 3; Length 1762;
Best Local Similarity 98.3%; Pred. No. 3.9e-281;
Matches 1043; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Db 625 AACTATACATCTGGGGGATCCATCACCAGCAGCAAAACAANTGAGCAGCAAAATTTGTATA 684
QY
Db 661 TCCAAAGATCAGGACGAGTAACAGTCTCAACAAAAAGAGTCAACAAACGATAGTCCCTA 720
Db 685 TCCAAAGATCAGGACGAGTAACAGTCTCAACAAAAAGAGTCAACAAACGATAGTCCCTA 744
QY 721 ATATCGGATCTAGACCGGTGGGTAGGGTCAATCAGGAGGATAGCATATATCTGGACCA 780
Db 745 ACATCGGATCTAGACCGGTGGGTAGGGTCAATCAGGAGGATAGCATATATCTGGACCA 804
QY 781 TTGTAAACCTTGGAGATATCTTAATGATAAAGTAAATGGCACTTAGTTGACCGGG 840
Db 805 TTGTAAACCTTGGAGATATCTTAATGATAAAGTAAATGGCACTTAGTTGACCGGG 864
QY 841 GATATTTAAATTTGAAACAGGAAAGCTCTGTAATGAGATCAGATCACCCTATAGACA 900
Db 865 GATATTTAAATTTGAAACAGGAAAGCTCTGTAATGAGATCAGATCACCCTATAGACA 924
QY 901 TTGTGTCTGTAATGTAATTTACACCAATGGAAGCATCCCAACGACAAACCATTTCAA 960
Db 925 CTGTGTCTGTAATGTAATTTACACCAATGGAAGCATCCCAACGACAAACCATTTCAA 984
QY 961 ATGTGAACAAATGTAATTTACACCAATGGAAGCATCCCAACGACAAACCATTTCAA 1020
Db 985 ATGTGAACAAATGTAATTTACACCAATGGAAGCATCCCAACGATATCAAGCAGAACTTTGAAGC 1044
QY 1021 TGGCCACTGGGATGAGGAATATACCAGAAAGCAAAATCAGA 1061
Db 1045 TGGCCACTGGGATGAGGAATATACCAGAAAGCAAAATCAGA 1085

RESULT 7

AAQ04597

ID AAQ04597 standard; DNA; 1797 BP.

XX AC AAQ04597;

XX 25-MAR-2003 (revised)

DT 02-OCT-1990 (first entry)

XX XX Equine hemagglutinin H3 (EIV-A2).

DE Recombinant vaccines; equine influenza virus; haemagglutinin; H3;

XX KW neuraminidase; N8; ss.

XX OS Equine influenza virus.

XX FH Key Location/Qualifiers

FT CDS 54..1748

FT /*tag= a

FT /product= "EIV-A2 haemagglutinin H3"

FT misc_RNA 102..1182

FT /*tag= b

FT /product= "N-terminal HA1 50kD portion"

FT misc_RNA 1186..1748

FT /*tag= c

FT /product= "C-terminal HA2 27kD portion"

XX US4920213-A.

XX 24-APR-1990.

XX 21-JUL-1986; 86US-00888250.

XX 20-JUN-1985; 85US-00747020.

XX (BIOT-) BIOTECH RES PARTNERS LTD.

XX Dale B, Cordell B;

XX WPI; 1990-163647/21.

XX P-PSDB; AAR04943.

XX

PT Recombinant vaccines against equine influenza virus - produced using DNA
PT sequences encoding haemagglutinin and neuraminidase glyco-protein(s).
XX Disclosure; Page ?; 27pp; English.

CC There are nine neuraminidase (NA) subtypes and twelve haemagglutinin (HA)
CC subtypes. The strain carrying H3N8 glycoproteins is designated equine
CC influenza virus (EIV)-A2. The cDNA sequences encoding these will be
CC useful in the construction of diagnostic probes for the disease and of
CC probes for obtaining new cDNAs of the mutated form of the virus.
CC Recombinant vaccines are produced. See also AAQ04596-Q04599. (Updated on
CC 25-MAR-2003 to correct PA field.)

XX SQ Sequence 1797 BP; 639 A; 341 C; 413 G; 404 T; 0 U; 0 Other;

Query Match 93.1%; Score 987.4; DB 2; Length 1797;

Best Local Similarity 95.7%; Pred. No. 1.8e-268;

Matches 1015; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 1 AGCAAAAGCAGGGGATATTTCTGTCAATCATGAGACAAACCATTTATTTGATACTACTGA 60
Db 25 AGCAAAAGCAGGGGATATTTCTGTCAATCATGAGACAAACCATTTATTTGATACTACTGA 84
QY 61 CCCATTGGGTCTACAGTCAAAACCCCAACCCAGTGGCAACACACAGCCACTATGTCTGG 120
Db 85 CCCATTGGGTCTACAGTCAAAACCCCAACCCAGTGGCAACACACAGCCACTATGTCTGG 144
QY 121 GACACCATGAGTAGCAAAATGGAAACATTTGTAATAAACAATACTGATGACCAATTTGAGG 180
Db 145 GACACCATGAGTAGCAAAATGGAAACATTTGTAATAAACAATACTGATGACCAATTTGAGG 204
QY 181 TGACAAATGCTACTGAAATTTAGTTCAGAGCACTTCAATAGATGCAATCTAGGAGACCCCACT 240
Db 205 TGACAAATGCTACTGAAATTTAGTTCAGAGCACTTCAATAGGAAATATGCAACCAACCAT 264
QY 241 ATAAAGTTCTAGATGGAAGAAATTTGCACATTAATAGATGCAATCTAGGAGACCCCACT 300
Db 265 ATAGGTTCTAGATGGAAGAAATTTGCACATTAATAGATGCAATCTAGGAGATCCCACT 324
QY 301 GTGATGCTTCCAGTATGAGAAATGGGACCTCTTCATAGAAAGAGCAGCGCTTTTCAGCA 360
Db 325 GTGATGCTTTCAGTATGAGAAATTTGGGACCTCTTCATAGAAAGAGCAGCGCTTTTCAGCA 384
QY 361 ATTGCTACCATATGACATCCCTGACATATGATCGCTCCGCTCCATTTAGTACCTCTCAG 420
Db 385 ATTGCTACCATATGACATCCCTGACATATGATCGCTCCGCTCTATTTGTGTCATCTTCAG 444
QY 421 GAACATTTAGAAATTCACAGCAGAGGATTCACATGGACAGCTGTCTACCTCAAAACCGAAGAA 480
Db 445 GAACATTTAGAAATTCACAGCAGAGGATTCACATGGACAGCTGTCTACCTCAAAACCGAAGAA 504
QY 481 GTGAGCCTGCAAAAGGGGATCAGCCGATAGTTCTTTTACCGGACTGAATTTGGCTAACAA 540
Db 505 GTGAGCCTGCAAAAGGGGATCAGCCGATAGTTCTTTTACCGGACTGAATTTGGCTAACAA 564
QY 541 AATCTGGAATCTTACCCCACTTGAATGTGAAATGTAACAAATGCTTAACTTTCGATA 600
Db 565 AATCTGGAATCTTACCCCACTTGAATGTGAAATGTAACAAATGCTTAACTTTCGATA 624
QY 601 AACTATACATCTGGGGGATTCATCACCAGCTCAAAACCAACAGCAGCAAGAAATTTGTACA 660
Db 625 AACTATACATCTGGGGGATTCATCACCAGCTCAAAACCAACAGCAGCAAGAAATTTGTACA 684
QY 661 TCCAAAGATCAGGACGAGTAAACAGTCTCAACAAAAAGAGTCAACAAACGATAGTCCCTA 720
Db 685 TCCAAAGATCAGGACGAGTAAACAGTCTCAACAAAAAGAGTCAACAAACGATAGTCCCTA 744
QY 721 ATATCGGATCTAGACCGGTGGGTAGGGTCAATCAGGAGGATAGCATATATCTGGACCA 780
Db 745 ACATCGGATCTAGACCGGTGGGTAGGGTCAATCAGGAGGATAGCATATATCTGGACCA 804
QY 781 TTGTAAACCTTGGAGATATCTTAATGATAAAGTAAATGGCACTTAGTTGACCGGG 840

Db 805 TTGTGAAACCTGGAGATATCTTAATGATAAACAGTAATGCGAACTTAGTTGCACCGGG 864
Qy 841 GATATTTTAAATTTGAAACAGGGAAGCTCTGTAATGAGATCAGATGCACCCATAGACA 900
Db 865 GATATTTTAAATTTGCGACAGGGAAGCTCTGTAATGAGATCAGATGCACCCATAGACA 924
Qy 901 TTTGTGTGCTGAATGTATTACCAACCAATGGAAGCATCCCCAAGCATATATCAGGCAAAACATTTTCAA 960
Db 925 CTTGTGTGCTCGAGTGTATTACCAACCAATGGAAGCATCCCCAAGCATATATCAGGCAAAACATTTTCAA 984
Qy 961 ATGTGAACAAAGTTACATATGGAATAATGCCCAAGTATATCAGGCAAAACATTTTAAAGC 1020
Db 985 ATGTGAACAAAGTTACATATGGAATAATGCCCAAGTATATCAGGCAAAACATTTTAAAGC 1044
Qy 1021 TGGCCACTGGATGAGGAATATACCGAAGCAAAATCAGA 1061
Db 1045 TGGCCACTGGATGAGGAATGTACCGAAGCAAAATCAGA 1085

RESULT 8

AAZ30211
ID AAZ30211 standard; DNA; 1698 BP.

XX AAZ30211;

XX 11-FEB-2000 (first entry)

XX Sequence of the haemagglutinin (HA) gene of EIV strain Newmarket 2/93.

XX Haemagglutinin gene; EIV, strain Newmarket 2/93; DNA vaccine; horse;
KW acrylic acid polymer; methacrylic acid polymer; copolymer;
KW maleic anhydride; alkenyl derivative; animal vaccine; viral infection;
KW bacterial infection; SS.

XX Equine influenza virus.

XX FR2776928-A1.

XX 08-OCT-1999.

XX 03-APR-1998; 98FR-00004409.

XX 03-APR-1998; 98FR-00004409.

XX (MERI-) MERIAL SAS.

XX Audonnet JCF, Minke JM;

XX WPI; 1999-593389/51.

XX Vaccine containing naked DNA and acrylic acid polymer or maleic anhydride
PT copolymer, for protection against viral or bacterial diseases in animals.

XX Example 8; Fig 1; 34pp; French.

XX The present sequence represents the haemagglutinin gene of Equine
CC influenza virus (EIV) strain Newmarket 2/93. The sequence was used to
CC prepare a DNA vaccine for horses, representative of the DNA vaccines of
CC the invention. The specification describes a DNA vaccine that comprises
CC naked DNA encoding an antigenic polypeptide, and at least one adjuvant
CC that is an acrylic or methacrylic acid polymer or a copolymer of maleic
CC anhydride with an alkenyl derivative. The vaccines are simple and easy to
CC prepare (simply by mixing components) and they do not involve any strong
CC interactions between DNA and other components that are likely to cause
CC complex formation. The vaccines are used to protect animals (pigs,
CC horses, dogs, cattle, cats or birds) against a wide variety of viral or
CC bacterial infections

XX Sequence 1698 BP; 619 A; 312 C; 372 G; 395 T; 0 U; 0 Other;

XX Query Match 92.78; Score 984; DB 2; Length 1698;

XX Best Local Similarity 97.14; Pred. No. 1.6e-267;

XX Matches 1002; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 30 ATGAAGACCAACCATTTATTTTGATCTACTGACCCATTGGGTCTTACAGTCAAAACCCCAACC 89
Db 1 ATGAAGACCAACCATTTATTTTGATCTACTGACCCATTGGGTCTTACAGTCAAAACCCCAACC 60
Qy 90 AGTGGAAAACAACAACGCCCATTTATGTCTGGGACACCATGCGAGTAGCAAAATGGAAACATTG 149
Db 61 AGTGGCAACAACAACGCCCATTTATGTCTGGGACACCATGCGAGTAGCAAAATGGAAACATTG 120
Qy 150 GTAAAAACAATAACTGATGACCAAAATGAGGTGACAAATGCTACTGTAATTTAGTTTCAGAGC 209
Db 121 GTAAAAACAATAACTGATGACCAAAATGAGGTGACAAATGCTACTGTAATTTAGTTTCAGAGC 180
Qy 210 ATTTCAATAGGAAAAATATGCAACCACTCATATAAAGTTCTAGATGGAAGAAATTCACACA 269
Db 181 ATTTCAATAGGAAAAATATGCAACCACTCATATAAAGTTCTAGATGGAAGAAATTCACACA 240
Qy 270 TTAATAGATGCAATGCTAGGAGACCCCACTGTGTGATGCTCTCCAGTATGAGAAATTTGGGAC 329
Db 241 TTAATAGATGCAATGCTAGGAGACCCCACTGTGTGATGATTTTCAGTATGAGAATTTGGGAC 300
Qy 330 CTCTTCATAGAAAGAGCAGCGCTTTACAGCAATTTGCTACCCATATGACATCCCTGACTAT 389
Db 301 CTCTTCATAGAAAGAGCAGCGCTTTACAGCAATTTGCTACCCATATGACATCCCTGACTAT 360
Qy 390 GCATCGCTCCGGTCCATTGTCATCCTCAGGAACATTTAGAAATTTACACAGAGGATTC 449
Db 361 GCATCGCTCCGGTCCATTGTCATCCTCAGGAACATTTAGAAATTTACACAGAGGAGTTTC 420
Qy 450 ACATGAGACAGGTGTCTACTCAAAACGGAAGCTGGAGCCTGCAAAAGGGGATCAGCCGAT 509
Db 421 ACATGAGACAGGTGTCTACTCAAAACGGAAGCTGGAGCCTGCAAAAGGGGATCAGCCGAT 480
Qy 510 AGTTTCTTTAGCCGACTGAATTTGGCTAAACAAAATCTGGAACCTTCTACCCCATTTGAAT 569
Db 481 AGTTTCTTTAGCCGACTGAATTTGGCTAAACAAAATCTGGAACCTTCTACCCCATTTGAAT 540
Qy 570 GTGACAAATGCCTAAACAAATTTTCGACAAACTATACATCTGGGGGATTTCAATACCCCG 629
Db 541 GTGACAAATGCCTAAACAAATTTTCGATAACTATACATCTGGGGGATTTCAATACCCCG 600
Qy 630 AGCTCAAAACCAACAGCAGACAGAAATTTGTATCCCAAGAAATCAGGACGAGTAAACAGTCTCA 689
Db 601 AGCTCAAAACCAACAGCAGACAGAAATTTGTATATCCCAAGAAATCAGGACGAGTAAACAGTCTCA 660
Qy 690 ACAAAAGAGTCAACAAACGATAGTCCCTTAATATCGGATCTAGACCGTGGGTTCAGGGT 749
Db 661 ACAAAAGAGTCAACAAACGATAGTCCCTTAATATCGGATCTAGACCGTGGGTTCAGGGT 720
Qy 750 CAATCAGGCGGATAGCATATATCTGGACCATTTGTAACACCTGGAGATATCCTTAATGATA 809
Db 721 CAATCAGGCGGATAGCATATATCTGGACCATTTGTAACACCTGGAGATATCCTTAATGATA 780
Qy 810 AACAGTAAATGGCAACATTAGTTGCAACCGGGGATATTTTAAATTTGAAAAACAGGAAAAAGC 869
Db 781 AACAGTAAATGGCAACATTAGTTGCAACCGGGGATATTTTAAATTTGAAAAACAGGAAAAAGC 840
Qy 870 TCTGTAATGAGATCAGATGACCCCATAGACATTTGTCTGCTGAATGTATTACACCAAT 929
Db 841 TCTGTAATGAGATCAGATGACCTCATAGACATTTGTGTCTGAATGTATTACACCAAT 900
Qy 930 GGAAGCATCCCCAACGACAAACCATTTTCAAAATGTGAACAAAGTTACATATGAAAAATGC 989
Db 901 GGAAGCATCCCCAACGACAAACCATTTTCAAAATGTGAACAAATTTACATATGAAAAATGC 960
Qy 990 CCCAAGTATATCGGCAAAACACTTTTAAAGCTGGCCACTGGGATGAGGAATATACAGAA 1049
Db 961 CCCAAGTATATCGGCAAAACACTTTTAAAGCTGGCCACTGGGATGAGGAATATACAGAA 1020
Qy 1050 AAGCAAAATCAGA 1061
Db 1021 AAGCAAAATCAGA 1032

```

RESULT 9
AAZ47007
ID AAZ47007 standard; DNA; 1698 BP.
XX
AC
XX
XX
DT 29-FEB-2000 (first entry)
XX
DE Equine influenza virus strain Newmarket 2/93 HA gene.
XX
KW Antibacterial; antiviral; primer; RT-PCR; amplification; haemagglutinin;
KW recombinant; vaccine; viral vector; pathogen; adjuvant; methacrylic acid;
KW maleic anhydride; alkenyl derivative; animal; herpes virus; tetanus;
KW influenza virus; feline leukemia; canine distemper; ss.
XX
OS Equine influenza virus.
XX
XX WO9944633-A1.
XX
XX 10-SEP-1999.
XX
XX 01-MAR-1999; 99WO-FR000453.
XX
XX 03-MAR-1998; 98FR-00002800.
XX
XX (MERI-) MERIAL.
XX
XX Audonnet JF, Minke JM;
XX
XX WPI; 2000-022918/02.
XX
XX Live recombinant vaccine comprising viral vector and polymeric adjuvant,
XX particularly directed against animal herpes and influenza viruses.
XX
XX Example 5; Fig 1; 41pp; French.
XX
XX This sequence represents the haemagglutinin (HA) gene from the equine
XX influenza virus strain Newmarket 2/93. The gene was amplified and the
XX product was used to generate a live recombinant vaccine which comprises:
XX (1) a viral vector including, and expressing in vivo, a heterologous
XX nucleotide sequence particularly a gene from a pathogen; and (2) at least
XX one adjuvant, i.e. a (meth)acrylic acid polymer or a copolymer of maleic
XX anhydride and alkenyl derivatives. The vaccines are used particularly to
XX protect against animal herpes or influenza viruses, but also feline
XX leukemia, tetanus and canine distemper
XX
XX Sequence 1698 BP; 619 A; 312 C; 372 G; 395 T; 0 U; 0 Other;
XX
XX Query Match 92.7%; Score 984; DB 3; Length 1698;
XX Best Local Similarity 97.1%; Pred. No. 1.6e-267;
XX Matches 1002; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
XX
QY 30 ATGAAGCAACCAATTTTGTATCTACTACTGACCCATTGGGTCTACAGTCAAAACCCCAACC 89
DB 1 ATGAAGCAACCAATTTTGTATCTACTACTGACCCATTGGGTCTACAGTCAAAACCCCAACC 60
QY 90 AGTGGAAACACACAGGCCCAATTTGTCTGGGACCACTGCTAGCAATGGAACATTTG 149
DB 61 AGTGGCAACACACAGGCCCAATTTGTCTGGGACCACTGCTAGCAATGGAACATTTG 120
QY 150 GTAAAAACAATACTGATGACCAAAATTTGAGTGCACAAATGCTACTGAATTTAGTTCAGAC 209
DB 121 GTAAAAACAATACTGATGACCAAAATTTGAGTGCACAAATGCTACTGAATTTAGTTCAGAC 180
QY 210 ATTTCAATAGGAAAAATATGCAACAACTCATATAAAGTTCTAGATGGAAGAAATTCACA 269
DB 181 ATTTCAATAGGAAAAATATGCAACAACTCATATAAGGTTCTAGATGGAAGAAATTCACA 240
QY 270 TTAATAGATGCAATGTAGAGAGACCCCACTGTGTCTTCCAGTATGAGAAATTTGGAC 329
DB 241 TTAATAGATGCAATGTAGAGAGACCCCACTGTGTGTGATGATTTTTCAGTATGAGAAATTTGGAC 300

```

```

QY 330 CTCTTCATAGAAAGACGCGCTTTTCAGCAATTTGCTACCATATGACATCCCTGACTAT 389
DB 301 CTCTTCATAGAAAGACGCGCTTTTCAGCAATTTGCTACCATATGACATCCCTGACTAT 360
QY 390 GCATCGCTCCGTCCTATTTAGCATCTCTAGGAACATTAGAAATTCACAGAGAGGATTC 449
DB 361 GCATCGCTCCGTCCTATTTAGCATCTCTAGGAACATTAGAAATTCACAGAGAGGATTC 420
QY 450 ACATGGACAGGTGTCCTCAAAACGGAAGTGGAGCCTGCAAAAGGGGATCAGCCGAT 509
DB 421 ACATGGACAGGTGTCCTCAAAACGGAAGTGGAGCCTGCAAAAGGGGATCAGCCGAT 480
QY 510 AGTTTCTTTAGCCGACTGAATTTGGCTAAACAAATCTGAAAACCTCTTACCCACACATTCGAT 569
DB 481 AGTTTCTTTAGCCGACTGAATTTGGCTAAACAAATCTGAAAACCTCTTACCCACATTCGAT 540
QY 570 GTGCAATGCTTAAACAAATTTTGGCAAACTATACATCTCTGGGGATTCATCACCCG 629
DB 541 GTGCAATGCTTAAACAAATTTTGGCAAACTATACATCTCTGGGGATTCATCACCCG 600
QY 630 AGCTCAAAACCAACAGCAGACAGAAATTTGTACATCCAGAAATCAGACAGATTAACAGTCTCA 689
DB 601 AGCTCAAAACCAACAGCAGACAGAAATTTGTACATCCAGAAATCAGACAGATTAACAGTCTCA 660
QY 690 ACAAAGAAAGTCAACAAACGATAGTCCCTTAAATATCGATCTAGACCGTGGGTAGGGGT 749
DB 661 ACAGAAAGAAAGTCAACAAACGATAGTCCCTTAAATATCGATCTAGACCGTGGGTAGGGGT 720
QY 750 CAATCAGGAGGATAAGCATATATCTGGAACATTTGTAAACCTGGAGATATCTCTAATGATA 809
DB 721 CAATCAGGAGGATAAGCATATATCTGGAACATTTGTAAACCTGGAGATATCTCTAATGATA 780
QY 810 AACAGTAATGGCAACTTAGTTGCAACCGCGGGATATTTTAAATTTGAAAACAGGAAAAGC 869
DB 781 AACAGTAATGGCAACTTAGTTGCAACCGCGGGATATTTTAAATTTGAAAACAGGAAAAGC 840
QY 870 TCTGTAATGAGATCAGATGCAACCATAGACATTTGTGTCTCAATGTATTACACCAAT 929
DB 841 TCTGTAATGAGATCAGATGCAACCATAGACATTTGTGTCTCAATGTATTACACCAAT 900
QY 930 GGAAGCATCCCCAACGACAAACCATTTTCAAAATGTGAACAAAGTTTACATATGAAAATGC 989
DB 901 GGAAGCATCCCCAACGACAAACCATTTTCAAAATGTGAACAAAGTTTACATATGAAAATGC 960
QY 990 CCCAAGTATATCAGGCAAAACACTTTTAAAGCTGGCCACTGGGATGAGGAATATACAGAA 1049
DB 961 CCCAAGTATATCAGGCAAAACACTTTTAAAGCTGGCCACTGGGATGAGGAATATACAGAA 1020
QY 1050 AAGCAAAATCAGA 1061
DB 1021 AAGCAAAATCAGA 1032
RESULT 10
AAQ29111
ID AAQ29111 standard; DNA; 1762 BP.
XX
AC AAQ29111;
XX
XX 25-MAR-2003 (revised)
XX 24-FEB-1993 (first entry)
XX
XX EIV HA (Al/Fontainbleau/79).
XX
XX Equine influenza virus; EIV; haemagglutinin; HA; Al/Fontainbleau/79;
XX expression cassette; NYVAC; ALVAC; recombinant vector;
XX polymerase chain reaction; PCR; vaccinia virus; H6 promoter;
XX canarypox virus; Copenhagen vaccine strain; virulence factor;
XX deletion loci; recipient loci; ss.
XX
XX Synthetic.
XX
XX WO9215672-A1.

```


XX PA (INMR) RHONE MERIEUX SA.
XX XX Audonnet JCF, Bouchardon A, Riviere MEA;
PI WPI; 1998-112826/11.
XX DR P-PSDB; AAW44946.
XX XX
PT Multi-valent polynucleotide vaccines against equine pathogens - consist
PT of at least 3 plasmids able to express protective antigens from specified
XX viruses.
XX
PS Example 14; Fig 8; 49pp; French.
XX XX
CC The invention relates to a multivalent vaccine for protecting horses
CC against several pathogens, especially pathogens associated with
CC respiratory and digestive diseases. The pathogens are especially selected
CC from equine herpesvirus (EHV), equine influenza virus (EIV), Clostridium
CC tetani, Borrelia burgdorferi, Eastern, Western or Venezuelan equine
CC encephalomyelitis viruses (EEV, WEEV and VEEV, respectively) and rabies
CC virus. The vaccines are preferably composed of polynucleotide sequences
CC encoding 3 antigens, all as part of vectors. This sequence represents the
CC coding region of the EIV Fontainebleau strain haemagglutinin gene. The
CC sequence was subcloned into the plasmid pVR1012 to generate plasmid
XX pAB099 for use in the vaccine
XX
SQ Sequence 1698 BP; 604 A; 319 C; 384 G; 391 T; 0 U; 0 Other;

Query Match 89.3%; Score 947.2; DB 2; Length 1698;
Best Local Similarity 94.9%; Pred. No. 4.2e-257;
Matches 979; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 30 ATGAAGACACCACTATTTTGTACTACTGACCCATTTGGTCTCTACAGTCAAAACCCAAC 89
DB 1 ATGAAGACACCACTATTTTGTACTACTGACCCATTTGGTCTCTACAGTCAAAACCCAAC 60
QY 90 AGTGGAAACACACAGCACCATATGTCTGGACACCATGAGTAGCAAAATGGAACATTG 149
DB 61 AGTGGCAACACACAGCACCATATGTCTGGACACCATGAGTAGCAAAATGGAACATTG 120
QY 150 GTAAAAACAATACTGTATGATGACCAAAATGAGGTGACCAAAATGCTACTGAATTTAGTTCAGAGC 209
DB 121 GTAAAAACAATACTGTATGATGACCAAAATGAGGTGACCAAAATGCTACTGAATTTAGTTCAGAGC 180
QY 210 ATTTCAATAGGAAATATCAACAACTCATATAAGTTCTAGATGCAAAATTTGACACA 269
DB 181 ACTTCAATAGGAAATATCAACAACTCATATAAGTTCTAGATGCAAAATTTGACACA 240
QY 270 TTAATAGATGCAATGTCTAGAGACCCCACTGTGATGCTTCCAGTATGAGAATTTGGAC 329
DB 241 TTAATAGATGCAATGTCTAGAGATCCCACTGTGATGCTTCCAGTATGAGAATTTGGAC 300
QY 330 CTCTTCATAGAAAGACAGCGCTTTTCAGCAATTTGCTACCCATATGACATCCCTGACTAT 389
DB 301 CTCTTCATAGAAAGACAGCGCTTTTCAGCAATTTGCTACCCATATGACATCCCTGACTAT 360
QY 390 GCATCGCTCCGTCATTTAGTATCTCTCAGGAACATTAGAAATTCACAGAGGATTC 449
DB 361 GCATCGCTCCGTCATTTAGTATCTCTCAGGAACATTAGAAATTCACAGAGGATTC 420
QY 450 ACATGGACAGTGTCTCTCAAAACGGAAGTGGAGCTGCAAAAGGGGATCAGCCGAT 509
DB 421 ACATGGACAGTGTCTCTCAAAACGGAAGTGGAGCTGCAAAAGGGGATCAGCCGAT 480
QY 510 AGTTTCTTTAGCCGACTGAATTTGGCTAAACAAATCTGGAACCTTTACCCCAATTTGAAT 569
DB 481 AGTTTCTTTAGCCGACTGAATTTGGCTAAACAAATCTGGAATTTTACCCCAATTTGAAT 540
QY 570 GTGCAATGCTTAACAATAAAATTTTCGACAACTATACATCTGGGGGATTCATCACCCG 629
DB 541 GTAAACATGCTTAACAATAAAATTTTCGATAAACTATACATCTGGGGGATTCATCACCCG 600
QY 630 AGCTCAAAACCAACAGCAGACAGAAATTTGATACATCCAGAAATCAGGACAGTAAACAGTCTCA 699

601 AGCACAACAAATGAGCAGACAAAATTTGTATGTCCTCAAGAAATAGGCGGAGTAAACAGTCTCA 660
QY 690 ACAAAAAGAGTCAACAAACGATAGTCCCTTAATATCGGATCTAGACCGTGGTGGGT 749
DB 661 ACAAAAAGAGTCAACAAACGATAGTCCCTTAATATCGGATCTAGACCGGCGGTGAGGGGT 720
QY 750 CAATCAGGAGGAGTAAAGCATATATCTGACCATTTGTAACCTGAGAGATATCTTAATGATA 809
DB 721 CAATCAGGAGGAGTAAAGCATATATCTGACCATTTGTAACCTGAGAGATATCTTAATGATA 780
QY 810 AACAGTAATGGCAACTTAGTTGACCCCGGGGATATTTTAAATTTGAAAAACAGGAAAAAGC 869
DB 781 AACAGTAATGGCAACTTAGTTGACCCCGGGGATATTTTAAATTTGAAAAACAGGAAAAAGC 840
QY 870 TCTGTAATGAGATCAGATGACCATCATAGATTTGTGTCTGTAATGTATTACACCAAT 929
DB 841 TCTTAATGAGATCAGATGACCATCATAGATTTGTGTCTGTAATGTATTACACCAAT 900
QY 930 GGAAGCATCCCCAACACCAACCATTTTCAAAATGTGAACAAAGTTACATATGAAAAATGC 989
DB 901 GGAAGCATCCCCAACACCAACCATTTTCAAAATGTGAACAAAGTTACATATGAAAAATGC 960
QY 990 CCCAAGTATATCAGGCAAAACACTTTTAAAGCTGGCCACTGGGATGAGGAATATACCAGAA 1049
DB 961 CCCAAGTATATCAGGCAAAACACTTTTAAAGCTGGCCACTGGGATGAGGAATATACCAGAA 1020
QY 1050 AAGCAAAATCAGA 1061
DB 1021 AAGCAAAATCAGA 1032

RESULT 12
AAQ72855
ID AAQ72855 standard; cDNA; 1777 BP.
XX
AC AAQ72855;
XX
DT 25-MAR-2003 (revised)
DT 23-JUN-1995 (first entry)
XX
XX Full length H3N2 influenza A virus, strain A2/Aichi/2/68 HA gene.
XX
KW Conserved peptide; stem region; hemagglutinin; HA; H1N1; H2N2; PCR;
KW subtype; human; influenza A virus; immunogenic artificial peptide;
KW antigen; vaccine; infection; polymerase chain reaction; primer; amplify;
KW C179; region A; region B; ds.
XX
OS Influenza A virus.
XX
FH Key Location/Qualifiers
CDS 36..1737
FT /*tag= a
FT /product= "Full length HA from H3N2 subtype"
FT sig_peptide 37..84
FT /*tag= b
FT misc_feature 85..246
FT /product= c
FT /product= "Stem region of N-terminal domain"
FT FT 247..903
FT /*tag= d
FT /product= "Globular head domain"
FT FT 904..1734
FT /*tag= e
FT /product= "Stem region of C-terminal domain"
XX
XX BP621339-A2.
XX
XX 26-OCT-1994.
XX
XX 20-APR-1994; 94BP-00302819.
XX
XX 20-APR-1993; 93JP-00115216.
PR

PR 16-MAR-1994; 94JP-00070194.
 XX (TAKI) TAKARA SHUZO CO LTD.
 PA
 PI Okuno Y, Isegawa Y, Sasao F, Ueda S;
 XX WPI; 1994-325949/41.
 DR P-PSDB; AAR63590.
 XX
 PT Human influenza-A virus haemagglutinin polypeptide(s) - useful in
 PT influenza-A vaccine composition.
 XX
 PS Example 2; Page 54-58; 68pp; English.
 XX
 CC This sequence was amplified using the primer sequences given in AAQ72852-
 CC 54 and represents a cDNA which encodes the full length hemagglutinin (HA)
 CC gene of the H3N2 subtype of human influenza A virus, strain
 CC A2/Aichi/2/68. The full length protein encoded by this amplified cDNA
 CC contains two conserved regions, the A' region, TGRN and the B region,
 CC QINGLINR(L/V)IEK. These regions are close to each other in the stem of
 CC the HA molecule and they represent epitopes which are recognised by the
 CC antibody C179. C179 binds to the stem region of the HA molecule and thus
 CC inhibits the membrane fusion action of the HA molecule and neutralises
 CC the virus. Polypeptide molecules which contain the conserved peptide
 CC regions, A and B, esp. HA molecules lacking the globular head region (see
 CC also AAQ72807), are antigenically equivalent to the stem region of the HA
 CC molecule of influenza A virus. These artificial peptides may be used as
 CC vaccines for prophylaxis of influenza A virus infection. (Updated on 25-
 CC MAR-2003 to correct PN field.)
 XX
 SQ Sequence 1777 BP; 565 A; 363 C; 420 G; 429 T; 0 U; 0 Other;

Query Match 57.2%; Score 607.2; DB 2; Length 1777;
 Best Local Similarity 74.0%; Pred. No. 6.3e-161;
 Matches 784; Conservative 0; Mismatches 273; Indels 3; Gaps 1;

QY 5 AAAGCAGGGGATATTTCTGTCAATCATGAAGACCAACCATTTATTTTATCTACTGACCCCA 64
 DB 12 AAAGCAGGGGATATTTCTATTAATCATGAAGACCATTTGCTTTGAGCTATATTTCTG 71
 QY 65 TTGGGCTTACAGTCAAAAC---CAAACCACTGGAAACAAACACAGCCACATTTATGTCGGG 121
 DB 72 TCTGGCTCTCGGCCAAGACCTTCCAGGAATGACACAGCAGCAGCGTGTGCTGGG 131
 QY 122 ACACCATGCTAGCAATGGAACATTTGGTAAACCAATCACTGATGACCAATTTAGGT 181
 DB 132 ACATCATGCGGTGCCAAACGGAACACTAGTGAACCAATCACAGATGATCAGATTGAAGT 191
 QY 182 GACAAATGCTACTGAATTAGTTTCAGAGCATTTCAATAGGGAATAATGCAACAACTCATA 241
 DB 192 GACTAATGCTACTGAGCTAGTTTACAGCTCTCAACGGGGAATAATGCAACAACTCTCA 251
 QY 242 TAAAGTTCTAGATGGAGAAATTCACATTAATAGATGCAATCTAGGAGACCCCACTG 301
 DB 252 TCGAATCCTTGTATGGAATAGACTGCACATGATAGATGCTTATTTGGGGACCTCATTTG 311
 QY 302 TGAATGCTTCAGTATGAGAAATTTGGGACCTTCTATAGAAAGACGAGCGCTTTCAGCAA 361
 DB 312 TGAATGTTTTTCAAAATGAGACATGGGACCTTTTCGTTGAACGAGCAAAAGCTTTTCAGCA 371
 QY 362 TTGCTACCCATATGACATCCTGACTATGATGATCGCTCCGCTCCATTTGATGATCCTCAGG 421
 DB 372 CTGTTACCTTATGATGTCAGATTAAGCTCCCTTAGTCTACTAGTTGCTCTGTCAGG 431
 QY 422 AACATTAGAATTCACAGCAGAGGGAATTCACATGACAGAGGTGTCACTCAAAACGGAAGAAG 481
 DB 432 CACTCTGGAGTTTATCACTCAGGGTTTCACTTTGAGCTGGGGTCACTCAGAATGGGGAAG 491
 QY 482 TGGAGCCTGAAAGGGGATCAGCCGATAGTTTCTTTAGCCGACTGAATTTGGCTAACAAA 541
 DB 492 CAATGCTTGCAAAAGGGGACCTGGTAGCGGTTTTTTCAGTAGACTGAACCTGGTTGACCAA 551
 QY 542 ATCTGGAACCTCTTACCCCAATTTGAATGTGACAAATGCCTTAACAAATAAAATTTTCGACAA 601

DB 552 ATCAGGAGCACATATCTCAGTCTGAACTGACTATGCAACAAATGACAAATTTTGACAA 611
 QY 602 ACTATACATCTGGGGATTTCACCCGAGCTCAACCAACAGCAGACAGAAATTTGACAT 661
 DB 612 ACTATACATTTGGGGGATTTCACCCGAGCAGAAACCAAGAACCAACCCAGCCTGTATGT 671
 QY 662 CCAAGNAATCAGGACGAGTAACAGTCTCAACAAAGAGAGTCAACAAACGATAGTCCCTAA 721
 DB 672 TCAAGCATCAGGAGAGTACAGTCTCTACAGGAGAGCCAGCAAACTATATATCCCGAA 731
 QY 722 TATCGGATCTAGACCGTGGTTAGGGGTCAATCAGCAGGATAAGCATATCTATGGAACAT 781
 DB 732 TATCGGTCAGACCCCTGGGTAAGGGTCTGTCTAGTAAGATTAAGCATCTATTGGACAAT 791
 QY 782 TGTAAACCTCGAGATATCTTAATGATAAACAAGTAATGGCAACTTAGTTGGACCGGGGG 841
 DB 792 AGTTAAGCCGGGAGAGCTACTGCTAATTAATAGTAATGGGAACCTAATCGCTCCTCGGG 851
 QY 842 ATATTTTAAATTTGAAACACAGGGAAGCTCTCTAATGAGATCAGATGCAACCCATAGCAT 901
 DB 852 TTATTTCAAAATGCGCACCTGGGAAAGCTCAATAATGAGTCAGATGCACTTATTGATAC 911
 QY 902 TTGTGTCTGAATGATTATACACCAATGGAAGCATCCCAACGACAAACCAATTTTCAAAA 961
 DB 912 CTGTATTTCTGAATGATCACTCCAAATGGAAGCATTTCCCAATGACAGCCCTTTCAAA 971
 QY 962 TGTGAACAAAGTTACATATGGAATAATGCCCAAGTATATCAGGCAAAACACTTTAAAGCT 1021
 DB 972 CGTAAACAAAGATCACATATGGAGCATGCCCAAGTATGTTAAGCAAAACACCCCTGAAGT 1031
 QY 1022 GCCCACTGGGATGAGGAATATACCAAGAAAGCAAAATCAGA 1061
 DB 1032 GGCAACAGGGATGCGGAATGTACCAAGAGAAACAAACTAGA 1071

RESULT 13
 ID ABA93937 standard; DNA; 1765 BP.
 XX ABA93937;
 XX AC
 XX DT 07-MAY-2002 (first entry)
 XX DE Influenza A/Udorn/72 (H2N2) Strain HA encoding DNA SEQ ID NO:7.
 XX KW Influenza A/Udorn/72 (H2N2) strain; Influenzavirus A; diagnosis;
 XX KW Influenza A virus; genome; gene; ds.
 XX OS Influenzavirus A.
 XX FH Key Location/Qualifiers
 XX CDS 30..1730
 XX FT /*tag= a
 XX FT /product= "HA protein"
 XX PN WO200200884-A2.
 XX PD 03-JAN-2002.
 XX PF 21-JUN-2001; 2001WO-US019826.
 XX PR 23-JUN-2000; 2000US-0213650P.
 XX PA (AMCY) AMERICAN CYANAMID CO.
 XX PI Galarza JM, Latham TE;
 XX DR WPI; 2002-139923/18.
 XX DR P-PSDB; ABB05767.
 XX PT Polynucleotide encoding complete sequence of influenza A/Udorn/72 and
 PT polypeptide, useful in diagnosis and for generating new influenza A

variant strains.

Claim 1; Page 61-64; 103pp; English.

The present invention describes an isolated polynucleotide (I) having the complete sequence of the Influenza A/Udorn/72 (H3N2) strain in positive strand, antigenomic message sense. ABA93944 encode the Influenza A/Udorn/72 (H3N2) strain proteins given in ABB05764 to ABB05774 from the present invention. (I) is useful for designing polymerase chain reaction (PCR) primers for use in a PCR assay to detect the presence of the corresponding virus segment in a sample or for designing and selecting peptides for use in an enzyme linked immunosorbant assay to detect the presence of the corresponding protein produced by that segment in a sample, hence is useful in diagnosis and may be modified by mutation to generate new influenza A variant strains. ABA94945 to ABA94039 represent Influenza A/Udorn/72 (H3N2) strain sequencing primers, which are used in an example from the present invention

Sequence 1765 BP; 568 A; 355 C; 414 G; 428 T; 0 U; 0 Other;

Query Match 56.7%; Score 601.6; DB 6; Length 1765;
Best Local Similarity 73.5%; Pred. No. 2.4e-159;
Matches 782; Conservative 0; Mismatches 279; Indels 3; Gaps 1;

Qy 1 AGCAAAAGCAGGGGATATTTCTGTCAATCATGAAGACAAACCATATTTTGTACTACTGA 60
Db 1 AGCAAAAGCAGGGGATATTTCTGTCAATCATGAAGACAAACCATATTTTGTACTACTAT 60

Qy 61 CCCATTGGTCTACAGTCAAAAC---CCAAACAGTGGAAACACACAGGCCACATTTATGTC 117
Db 61 TCTGTCTGGTCTCGGCCAAACCTTCCAGGAATGACACAGCACAGCAACGCTGTGCC 120

Qy 118 TGGGACACCATGAGTACCAATGGAATGTTGTAACCAATGTAACATGATGACCAATTTG 177
Db 121 TGGGACATCATCGCGTCCCAACGACACTAGTGAACCAATGTAACATGATGATGATG 180

Qy 178 AGGTGACAAATGCTACTGTAATGCTCAGAGCATTTTCAATAGGGAAATATGCAACACT 237
Db 181 AAGTGACTAATGCTACTGAGCTGTTTCAAGTCTTCAACGGGAAATATGCAACATC 240

Qy 238 CATATAAAGTTCTAGATGGAAGAAATGCAATTAATAGATGCAATGCTAGGAGACCCCC 297
Db 241 CTCATCAATCTTGTAGTGAATAGCTGCACACTGATAGATGCTTCTATTGGGGACCCCT 300

Qy 298 ACTGTGATCTTCCAGTATGAGAAATGGGACCTTTCATAGAAGAGCAGCGCTTTCA 357
Db 301 ATTGTGATGGCTTTCAAAATGAGACATGGGACCTTTTCGTGAAACGACGACATTTCA 360

Qy 358 GCAATTTGCTACCCATATGACATCCCTGACTATGCTATGCTCCGCTCCATTGTAGCATCCT 417
Db 361 GCACTGTACCTTATGATGTGCGAGTATGCTTCCCTTAGGTCTACTAGTTGCTCTGT 420

Qy 418 CAGGAACATTAGAAATCAGCAGAGAGGATTCATCGACAGGTGTCTCACTCAAAACGGAA 477
Db 421 CAGGCACTCTGGAGTTTATCAGTGAAGGCTTCACTTGGACTGGGGTCACTCAGAATGGG 480

Qy 478 GAGTGGAGCTGCAAAAGGGATCAGCCGATGCTTCTTAGCCGACTGAATTTGGCTAA 537
Db 481 GAAGCAATGCTTCAAAAGGGACCTGATAGCGGTTTTTTTTCAGTAGACTGAATGTTGT 540

Qy 538 CAAATCTGGAACCTTTACCCCAATGTAATGTGACAAATGCTTAAACATAAAATTTTCG 597
Db 541 ACAANTCAGNAGCAGATATCCAGTGTGCTGAACGTAATGCTCCAAACATGACATTTTG 600

Qy 598 ACAACATATACATCTGGGGATTCATCCCGAGCTCAAAACCAACAGCAGACAGAAATTTG 657
Db 601 ACAACATATACATTTGGGGGTTTCAACACCGGACCGGACCAAGAACAAACCGACCTAT 660

Qy 658 ACATCCAGAAATCAGGACGAGTATACAGTCTCAACAAAAGAGTCAACAAACGATGCTCC 717
Db 661 ATGTTCAAGCATCAGGGAGAGTCACTCTCTTACCAAGAGAAGCGACCAAACTATATATCC 720

Qy 718 CTAATATCGGATCTAGACCGTGGGTTAGGGGTCAATCAGGAGATGAAGCATATACTGGA 777

Db 721 CGAATATCGGGTCTAGACCTGGGTAGGGGTCTGTCTAGTAGAATAAGCATCTATTGGA 780

Qy 778 CCATTGTAAACCTGGAGATATCTTAATGATAAAGCACTTAATGGCACTTAGTTGACCGC 837
Db 781 CAATAGTTAAACCGGGAGACATCTGCTAATTAATAGTAATGGGAACCTAATTTGCTCCTC 840

Qy 838 GGGGATATTTTAAATGAAAACAGGGGAAAGCTCTGTAATGAGATCAGATGCACCCATAG 897
Db 841 GGGGTTATTTTAAATGCGCACTGGGAAAGCTCAATAATGAGTCAAGTGCACCTATTG 900

Qy 898 ACATTTGTGTCTGTAATGTAATACCAATGGAAGCATCCCCAAGCAAAACCATTTTC 957
Db 901 GCACCTGCAATTTCTGTAATGATCATCTCCAAATGGAAGCAATCCCAATGACAGCCCTTC 960

Qy 958 AAAATGTGAACAAAGTTACATATGGAATAATGCCCAAGTATATCAGGCAAAACACTTTAA 1017
Db 961 AAAACGTAACAAAGATCACATATGGGTCATGTCCCAAGTATGTAAAGCAAAACACCTGA 1020

Qy 1018 AGCTGGCCACTGGGATGAGGAATATATACCAAGAAAGCAAAATCAGA 1061
Db 1021 AGTTGGCAACAGGATGCGGAATGTACCAAGAAACAAACTAGA 1064

RESULT 14
ABA93944
ID ABA93944 standard; DNA; 1764 BP.
XX ABA93944;
AC AC
XX XX
DT 07-MAY-2002 (first entry)
XX
DE Influenza A/Udorn/72 (H3N2) Strain HA encoding DNA SEQ ID NO:21.
XX
KW Influenza A/Udorn/72 (H3N2) strain; Influenzavirus A; diagnosis;
XX Influenza A virus; genome; gene; ds.
OS Influenzavirus A.
FH Key Location/Qualifiers
FT CDS 30..1730
FT /*tag= a
FT /product= "HA protein"
XX
PN WO200200884-A2.
XX
PD 03-JAN-2002.
XX
PF 21-JUN-2001; 2001WO-US019826.
XX
PR 23-JUN-2000; 2000US-0213650P.
XX
PA (AMCY) AMERICAN CYANAMID CO.
XX
PI Galarza JM, Latham TE;
XX
DR WPI; 2002-139923/18.
DR P-PSDB; ABB05774.
XX
PT Polynucleotide encoding complete sequence of Influenza A/Udorn/72 and
PT polypeptide, useful in diagnosis and for generating new influenza A
PT variant strains.
XX
PS Claim 1; Page 80-83; 103pp; English.
XX
CC The present invention describes an isolated polynucleotide (I) having the
CC complete sequence of the Influenza A/Udorn/72 (H3N2) strain in positive
CC strand, antigenomic message sense. ABA93944 to ABA93944 encode the
CC Influenza A/Udorn/72 (H3N2) strain proteins given in ABB05764 to ABB05774
CC from the present invention. (I) is useful for designing polymerase chain
CC reaction (PCR) primers for use in a PCR assay to detect the presence of
CC the corresponding virus segment in a sample or for designing and
CC selecting peptides for use in an enzyme linked immunosorbant assay to
CC detect the presence of the corresponding protein produced by that segment
CC in a sample, hence is useful in diagnosis and may be modified by mutation
CC to generate new influenza A variant strains. ABA94945 to ABA94039
CC represent Influenza A/Udorn/72 (H3N2) strain sequencing primers, which
CC are used in an example from the present invention

CC detect the presence of the corresponding protein produced by that segment
 CC in a sample, hence is useful in diagnosis and may be modified by mutation
 CC to generate new influenza A variant strains. ABA94945 to ABA94039
 CC represent Influenza A/Udmr/72 (H3N2) strain sequencing primers, which
 CC are used in an example from the present invention
 XX

SQ Sequence 1764 BP; 568 A; 355 C; 414 G; 427 T; 0 U; 0 Other;

*Query Match 56.4%; Score 598.4; DB 6; Length 1764;
 Best Local Similarity 73.3%; Pred. No. 1.9e-158;
 Matches 780; Conservative 0; Mismatches 281; Indels 3; Gaps 1;

QY 1 AGCAAAAGCAGGGGATATTTCTGTCATCATGAGCAACACCATTTTGTACTACTGA 60
 DB 1 AGCAAAAGCAGGGGATATTTCTGTCATCATGAGCAACACCATTTTGTACTACTGA 60
 QY 61 CCATTTGGGTCTACAGTCAAAAC---CCAAACAGTGGAAACAAACAGCAGCATTATGTC 117
 DB 61 TCTGTCTGTCTCGGCCAAGACTTTCCAGGAAATGACACAGCAGCAACGCTGTGCC 120
 QY 118 TGGGACACCATCAGTAGCAAAATGGACATTTGTTAAACAAATAAATGATGACCAATTTG 177
 DB 121 TGGGACATCATCGGTGCCAAACGGAACACTAGTGAACAAATCACAATATGATCAGATTG 180
 QY 178 AGGTGACAAATGCTTACTGAATATTAGTTTCAGAGCATTTTCAATAGGGAATAATGCAACA 237
 DB 181 AAGTGACTAATGCTTACTGAGCTGTTTCAGAGTTTCTTCAACGGGGAATAATGCAACA 240
 QY 238 CATATAAAGTCTAGATGGAAGAAATGACATTAATAGATGCAATGCTAGGAGACCC 297
 DB 241 CTATCGAATCCTTGATGGAATAGATGCAACACTGATGCTCTATTGGGGGACCCCT 300
 QY 298 ACTGTGATGCTTCCAGTATGAGAAATGGGACCTCTTCATAGAAAGAGCAGCGCTTCA 357
 DB 301 ATTGTGATGCTTTCAAAATGAGACATGGGACCTTTTCGTTGAAACGAGCAAGCTTTCA 360
 QY 358 GCAATTTGCTACCATATGACATCCCTGACTATGATGCTGCTCGGTCCATTTGTAGCATCCT 417
 DB 361 GCAACTGTTTACCCTTATGATGTCAGATTTATGCTTATGCTTACCTAGTCTGCTCGT 420
 QY 418 CAGGAACATTAGAAATTCACAGCAGAGGGATTCATGACGACAGTGTCTCAAAACGGA 477
 DB 421 CAGGCACTCTGAGATTTATCAGTGAAGGCTTCACTTGGACTGGGGTCACTCAGAAATGGG 480
 QY 478 GAAGTGGAGCTCGAAAGGGGATCAGCGATAGTTTCTTTAGCCGACTGAATTTGCTAA 537
 DB 481 GAAGCAATGCTTGAAGAGGGGACCTGATAGCGGTTTTCATGACTGAACTGTTGTT 540
 QY 538 CAAATCTGGAAACTCTTTACCCACATTTGAATGTGCAATGCTTCAACAAATFAAAATTTTCG 597
 DB 541 ACAATCAGGAAGCACATATCCAGTGTGAACTGACTATGCCAAACATGACAAATTTTG 600
 QY 598 ACAACATATACATCTGGGGATTCATCACCGAGCTCAACCAACAGCAGACAGAAATGT 657
 DB 601 ACAACATATACATTTGGGGGTTTCAACCCGAGCAGCGCAACGAACCAACAGCGCTAT 660
 QY 658 ACATCCAAGATCAGGACGAGTAACTCTCAACAAAGAAAGTCAACAAACATAGTCC 717
 DB 661 ATGTTTACGATCAGGGAGAGTCAAGTCTTACCAAGAGAGCAGCAACATATATATCC 720
 QY 718 CTAATATCGGATCTAGACCGTGGGTTAGGGGTCAATCAGGAGGATAAGCATATCTGGA 777
 DB 721 CGAATATCGGTTCTAGACCTCGGTAAGGGGTCTGTCTAGTAGAATAAGCATCTATTGGA 780
 QY 778 CCATTGTAAACCTGGAGATATCTTAATGATAACAGTAAATGCAACTTATGTTGACCCG 837
 DB 781 CAATAGTAAACCGGGAGACATCTGGTAAATTAATAGTAAATGGAACCTTAATTTGCTCCT 840
 QY 838 GGGGATATTTAAATGAAACAGGGGAAAGCTCTGTAAATGAGATCAGATGACACCCCATAG 897
 DB 841 GGGGTTATTTAAATGCGGACCTGGGAAAGCTCAATATAGAGTTCAGATGACCTATTG 900
 QY 898 ACATTTGTGTCTGTAATGATTTACACCAAAATGGAAGCATCCCCAAGCAACCAATTTTC 957

DB 901 GCACCTGTCATTTCTGAATGCATCACTCCAAATGGAAGCATTCCTCAATGACAAGCCCTTTC 960
 QY 958 AAAATGTGACAAAGCTTACATATGGAATGCCCCAAGTATATCAGGCCAAACACTTTAA 1017
 DB 961 AAAACGTAAACCAAGATCACAATATGGGGCATGTGCCAAGTATGTTAAGCAAAACACCTGA 1020
 QY 1018 AGCTGGCCACTGGGATGAGGAATATACACAGAAAGCAAAATCAGA 1061
 DB 1021 AGTTGGCAACAGGATGCGGAATGTACCAAGAAACAAACTAGA 1064

RESULT 15
 AAN70642
 ID AAN70642 standard; cDNA; 1762 BP.
 AC AAN70642;
 XX
 DT 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 29-APR-1991 (first entry)
 XX
 DE Sequence encoding influenza HA protein.
 XX HA; vaccine; ds.
 XX Influenza A virus; 'A/Niigata 102181'.
 XX JP62051992-A.
 XX
 PD 06-MAR-1987.
 XX
 XX 30-AUG-1985; 85JP-00192915.
 XX
 XX 30-AUG-1985; 85JP-00192915.
 XX
 XX (KAGA) KAGAKU KASEI RYOHO KENKYUSHO.
 XX WPI; 1987-104909/15.
 XX
 XX Recombinant plasmid - combines influenza virus gene and transformation enzyme and can be used for prepn. of influenza virus protein.
 XX Disclosure; Fig 3; 14pp; Japanese.
 XX
 CC Sequence may be inserted into a plasmid under the control of a phosphatase promoter, and used to transfer an E.coli expression system. The HA protein product may be used as vaccine against influenza viral infection. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 27-AUG-2003 to correct OS field.)
 XX
 SQ Sequence 1762 BP; 579 A; 359 C; 404 G; 420 T; 0 U; 0 Other;

Query Match 55.2%; Score 585.6; DB 1; Length 1762;
 Best Local Similarity 72.6%; Pred. No. 8e-155;
 Matches 772; Conservative 0; Mismatches 289; Indels 3; Gaps 1;
 QY 1 AGCAAAAGCAGGGGATATTTCTGTCATCATGAGCAACACCATTTTGTACTACTGA 60
 DB 1 AGCAAAAGCAGGGGATATTTCTGTCATCATGAGCAACACCATTTTGTACTACTGA 60
 QY 61 CCATTTGGGTCTACAGTCAAAAC---CCAAACAGTGGAAACAAACAGCAGCATTATGTC 117
 DB 61 TCTGTGAGTCTCGGCCAAGACTTTCCAGGAAATGACACAGCAGCAACGCTGTGCC 120
 QY 118 TGGGACACCATCAGTAGCAAAATGGAAACATTTGGTAAAAACAATACTGATGACCAATTTG 177
 DB 121 TGGGACATCATGTAGTGCACAAACGGAACGCTAGTGAACAAATCAGAAATGATCAGATTG 180
 QY 178 AGGTGACAAATGCTTACTGAAATTAGTTTCAGAGCATTTTCAATAGGGAATAATGCAACA 237
 DB 181 AAGTGACTAATGCTTACTGAGCTGGTTTCAGAGTTTCTCAACAGTAGAATATGCAAGATC 240

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

Db 93 AWAAMAMWMAAAAAAAMWMAWGAWANGWGWGAWMMWMAWNGMMWMAWMAAAAA 152

Qy 822 AACTTACTTGACCGCGGGATATTTTAAATTGAAAAACGGGAAAGCTCTGTAATGAGA 881

Db 153 WMMWCTTTGCMCTGAWTRTWGTRTWAWMMWTRTAAWMMWMAWMAWMAWMAWMAW 212

Qy 882 TCAGATGACCCATAGACATTTTGCTGCTGTAATGTAATACACAAATGGAAGCATCCCC 941

Db 213 WMMWMAAARAAMWMAWMMRCRWTWCRCMAWMAWMAAAMAAWMAWMAWMAWMAW 272

Qy 942 AACGACAAACATTTCAAAATGTGAACAAAGTTACATATGCAAAATGCCCAAGTATATC 1001

Db 273 MAAMWCSAAWGWGAWAAWTRGAWMMWMAWMAAAMRAARMWMAWMAWMAWMAWMA 332

Qy 1002 AGGCAAAACACTTTAAA 1018

Db 333 CAWMAWAAARTWRTA 349

RESULT 2
AO636729
LOCUS
DEFINITION
RPCI-11-477P23-TV RPCI-11 Homo sapiens genomic clone
RPCI-11-477P23, genomic survey sequence.

ACCESSION
AO636729
VERSION
AO636729.1 GI:5099364
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
1 (bases 1 to 598)
Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
Venter,J.C.

AUTHORS
Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
Venter,J.C.

TITLE
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready

JOURNAL
Map Building

COMMENT
Unpublished (1997)
Other GSSs: RPCI-11-477P23.TJ
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
7712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@igr.org

Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from
Research Genet cs (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.

Seq primer: T7
Class: BAC ends.

FEATURES
source
1..598
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7683166"
/db_xref="taxon:9606"
/clone="RPCI-11-477P23"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPCI-11"
/notes="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"

ORIGIN
Query Match 4.1%; Score 43.8; DB 9; Length 598;
Best Local Similarity 56.4%; Pred. No. 1.3;
Matches 101; Conservative 0; Mismatches 77; Indels 1; Gaps 1;

Qy 100 ACACAGCCACATTATGTCTGGGACACCATGCAAGTGAACATTGGTAAAAACAA 159

Db 198 ATATAGCAGAAACACATTTATGACACGATTACATTATACAGAAATGTTGACTAAAGAA 257

Qy 160 TAACTGATGACCAAAATGAGGTGACAAATGCTACTGAATTAGTTGAG-AGCATTTCATA 218

Db 258 TAACCAATTTAAAAAATTTGAAGAGTGAATTTTAAATTTAAATTCAGTAAAAAGTACAGA 317

Qy 219 GGGAAAAATATGCAACAACTCATATAAAGTTCTAGATGGAAGAAATGCGACATTAAATAGA 277

Db 318 TGTAGAATATGAACAAATTTATATTAATGATAAAGAAACATTGTCTTAATATA 376

RESULT 3
CC081372/c
LOCUS
DEFINITION
CSU-K33r.31E9.SP6 CSU-K33r Aedes aegypti genomic clone
CSU-K33r.31E9, genomic survey sequence.

ACCESSION
CC081372
VERSION
CC081372.1 GI:29929272
KEYWORDS
GSS.
SOURCE
Aedes aegypti (yellow fever mosquito)

ORGANISM
Aedes aegypti
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
Culicidae; Culicinae; Culicini; Aedes; Stegomyia.

REFERENCE
1 (bases 1 to 820)
Loftus,B., Shetty,J., Severson,D., Brown,S. and Knudson,D.
Aedes aegypti

AUTHORS
Loftus,B., Shetty,J., Severson,D., Brown,S. and Knudson,D.

TITLE
End sequencing of Aedes aegypti BACS

JOURNAL
Unpublished (2003)

COMMENT
Other GSSs: CSU-K33r.31E9.T7
Contact: Brendan Loftus
Department of Eukaryotic Genomics
TIGR

7712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208
Email: enta@tigr.org

Library was provided by Susan Brown and Dennis Knudson at Colorado
State University.
Seq primer: SP6
Class: BAC ends.

FEATURES
source
1..820
Location/Qualifiers
/organism="Aedes aegypti"
/mol_type="genomic DNA"
/strain="Rexville"
/db_xref="taxon:7159"
/clone="CSU-K33r.31E9"
/clone_lib="CSU-K33r"
/note="Vector: pBelobAC11; Site_1: HindIII"

ORIGIN
Query Match 4.1%; Score 43.8; DB 9; Length 820;
Best Local Similarity 49.0%; Pred. No. 1.4;
Matches 117; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

Qy 42 ATTATTTTGTACTACTGACCCATGGGTCTACAGTCAAACCCACCAGTGAACAAC 101

Db 284 ATCATTTCAAACCTGTTTTCGTTAGTTACACGGAAATAAATAACGCTCCGAAAAA 225

Qy 102 ACAGCCACATTTATGTCCTGGGACACCATGCAAGTGAACATTTGGTAAAAACAATA 161

Db 224 AAATAAAATTTAAATATTTTAAATAACCGAAACAATTTTATTTATTGCGAAA 165

Qy 162 ACTGATGACAAATTTGAGGTGCAAAATGCTACTGAATTTAGTTTCAGAGCATTTCAATAGG 221

Db 164 AATCAACAACACTAGAAAAGGCTTCAGAAAAAATGAAAAAGCTAGGGATTTCAAAAA 105

Qy 222 AAAATATGCAACAACCTCATATAAAGTTCTAGATGGAAGAAATTCGACATTAAATAGATGC 280

Db 104 AAAATTATAAAATCAAAAACCAAAATTTAAAAATTTTGAATATAAAATTAATAATGC 46


```

RESULT 4
CNS00905/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR19H07 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
VERSION
AL052980.1 GI:4934428
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 854)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1. .854
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19H07"
/clone_lib="RPCI-98"
/notes="end : T7"

FEATURES
source
ORIGIN
Query Match 4.0%; Score 42.8; DB 10; Length 854;
Best Local Similarity 30.3%; Pred. No. 2.6; Mismatches 358; Indels 3; Gaps 1;
Matches 158; Conservative 2;

Qy 536 AACAAAATCTGGAACCTTTACCCCAACATTCGATGTCACATGCTTACCAATATAAAATTT 595
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
626 AANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 567
Qy 596 CGACAAACTATACATCTGGGGATTCATCACCAGGCTCAACCAACAGCAGACAGAAAT 655
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
566 AANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 507
Qy 656 GTACATCAAGATCAGACAGGTAACAGTCTCAACAAAGAGTCAACAAACGATAGT 715
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
506 NNAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 447
Qy 716 CCCTTAATATCGATCTAGACCGTGGGTGATGAGGTCATCAGGCGAGGATAGCATATCTG 775
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
446 NNNNAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 390
Qy 776 GACCATTTGTAACCTGGAGATATCTTAATGATATAACAGTAATGGCACTTAGTTCACC 835
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
389 NNAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 330
Qy 836 GCGGGGATATTTAAATTTGAACAGGGAAGAGCTCTGTAATGAGATCAGATGCCCAT 895
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
329 NANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 270

```

```

896 AGACATTTGTGTCTGAATGTATTACACCAAAATGGAAGCATCCCAACGACAAACCAT 955
269 ANNNANATNNNNNAAGNNNTANNANNAANAANAANAANAANAANAANAANAANAANA 210
Qy 956 TCAAAATGTGAACAAAGTTACATATGGAATATGCCCAAGTATATCAGGCAAAACATTT 1015
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
209 NNNNNANNAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 150
Qy 1016 AAAGCTGCCACTGGGATGAGGATATACACGAAAGCAAA 1056
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
149 NNNNCADVKAACCAAAAGGAAAAAANNNNNNNAANAANAANAANAANAANA 109

RESULT 5
CNS0106X
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence T7 end of BAC
BACN03K20 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
VERSION
AL098595.1 GI:5610206
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.
Location/Qualifiers
1. .1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN03K20"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"
/notes="end : T7"

FEATURES
source
ORIGIN
Query Match 4.0%; Score 42.4; DB 10; Length 1101;
Best Local Similarity 19.3%; Pred. No. 3.7;
Matches 67; Conservative 140; Mismatches 141; Indels 0; Gaps 0;

Qy 563 ATTGAATGTGACAATGCCTAACAAATAAATTTTCGACAACTATACATCTGGGGATTC 622
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
754 RDKRRWDKDKKRRKAKAKAKAADADADAKADAKADAKADAKADADADGRGDGKKKRA 813
Qy 623 TCACCCGAGCTCAAAACCAACAGCAGACAGAAATTTGATACATCAAGAAATCAGCAGGTAAC 682
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
814 KDRKKKKKKDKAWGDKKKKAKKDKKAAKADAAAGDAKAKERRRAGDKDKADAKAKAK 873
Qy 683 AGTCTCAACAAAAGAGTCAACAAACGATAGTCCCTTAATATCGGATCTAGACCGTGGGT 742
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
874 AKDDDDAKAATKAKATKAKKDKAKAKKKKKDKDKAKAKAKADAKDKDDKDDKDK 933
Qy 743 TAGGGTCAATCAGGCGAGGATAGCATATCTAGCAACATTTGTAACCTGGAGATATCTCT 802
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
934 DAKADKKKKKKORAKKDKDKDKDKDKDKDKDKDKDKDKDKDKDKDKDKDKDKDK 993
Qy 803 AATGATAAACAAGTAATGGCAACTTAGTTGACCGCGGGGATATTTTAAATTTGAAACAGG 862

```


Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Akakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Iishi, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 12466851

Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216

Email: genome-resgsc.riken.jp, URL: <http://genome.gsc.riken.jp/>
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K., Iishi, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission
 Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN Integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

Location/Qualifiers
 1..665
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="B020021A19"
 /tissue_type="egg"
 /dev_stage="2 cells"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, 2 cells egg"
 /notes="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer 15'
 GAGAGAGAGAGGATCCAGAGCTCTTTTCTTTTCTTTT 3', cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence 15'
 GAGAGAGATTCGAGTGAATTAATTAATCCCCCCCCCC 3'. cDNA was cleaved with XhoI and BamHI. Vector: a modified

ORIGIN
 Query Match 4.0%; Score 42; DB 5; Length 665;
 Best Local Similarity 50.0%; Pred. No. 4.1;
 Matches 105; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
 pBluescript KS(+) after bulk excision from Lambda FLC I."

QY 819 GGCAACTTAGTTCGACCGGGGATATTTTAAATTTGAAACACGAGAAAGCTCTGTAATG 878
 |||||
 Db 177 GACAAACCAATGAATCTGTGTTCTATCTCATCTGTCGCCAGAGGTGAAACTCGACCCAT 236
 |||||
 QY 879 AGATCAGATGCACCCCATAGACATTTGTGTCTCTGAATGTATTACACCAATGGAAGCATC 938
 |||||
 Db 237 ATATCTCATCCACCTCGCAATAATTTGATTCAGATGGATCATCCGCCATTGATAGAAGA 296
 |||||
 QY 939 CCCAAGCAGCAACCAATTTCAAAATGTGAACAAAGTTACATATGGAATAATGCCCAAGTAT 998
 |||||
 Db 297 CCAGAAGCTCTGAAACTGCAGGATGAGAACAAATAAATGCTAGATAGATTTCATGAAA 356
 |||||
 QY 999 ATCAGGCAAAACACTTTAAAGCTGGCCACT 1028
 |||||
 Db 357 AGGACTCCAGGAATTTAGATGGTGGTCATT 386
 |||||

RESULT 8

LOCUS CNS00LRE/c 1106 bp DNA linear GSS 14-JUN-1999
 DEFINITION Drosophila melanogaster genome survey sequence 17 end of BAC:
 BAC48P05 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
 ACCESSION AL078654
 VERSION AL078654.1 GI:5101944
 KEYWORDS GSS.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1106)
 REFERENCE
 AUTHORS
 TITLE Direct Submission
 JOURNAL Submitted (11-JUN-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
 - Web : www.genoscope.cns.fr
 COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Ooeogawa and Aaron Hammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw ap, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or fillers for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
 1..1106
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BAC48P05"
 /clone_lib="RPCI-98"
 /note="end : T7"

source

ORIGIN

Query Match 3.9%; Score 41.6; DB 10; Length 1106;
 Best Local Similarity 15.4%; Pred. No. 6.1;
 Matches 55; Conservative 151; Mismatches 152; Indels 0; Gaps 0;

```
QY 588 AAAAAATTCACAAACTATACATCTGGGGATTTCATCACCCGAGCTCAACACCAACAGCAG 647
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1041 AHKDKRSNKKBGDSVNTVAVDNTBWSRRHNCNVRBDSCSVHWVTDYBCKNTYKWG 982
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 648 ACAGAAATTGATCCAGAAATCAGGACGAGTAAACAGTCTCAACAAAGAAAGAGTCAACAA 707
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 981 VSNKWBTKHYVMBCYKWBWHAHAAWAAWAAADAAACCCRCBRWVAVAWAAMRAMMAAAAA 922
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 708 ACATAGTCCCTAATATCGGATCTAGACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 767
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 921 MKHTSGBKYSYDMDHTTKYKWRWYKSGNGSGSHSGSGSMVAVWCNRYHAYTHRKWVWY 862
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 768 ATATCTGGACCATTTGTAACACCTGGAGATATCTCTAATGATAAACAAGTATGGCAACTTA 827
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 861 NWHBDSDBVCMCCVMAHWCHEBWHVCYYNANSCANBAWYVAKSBTRTHWNNYTCA 802
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 828 GTTGACCCGGGGATATTTTAAATTGAAACAGGAAAGCTCTGTAATGAGATCAGAT 887
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 801 ANGNTKWARGSGCACCBGMRYHTGGGRARGGKMGVCRCCVKBMWHAHHBNWHNNY 742
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 888 GCACCATAGACATTTGTGTCTGATGATGATTAACACCAATGGAAGCATCCCAACG 945
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 741 NNMAWAAVAMWNWATKYKKNACYKTKGVDVAYMRAVARCSRVRVNCVAVAVAWABG 684

RESULT 9
LOCUS A0055729/c 733 bp DNA linear GSS 30-JUL-1998
DEFINITION CIT-HSP-2341D11.TF CIT-HSP Homo sapiens genomic clone 2341D11,
genomic survey sequence.
ACCESSION A0055729
VERSION A0055729.1 GI:3352335
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 733)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
JOURNAL Unpublished (1998)
COMMENT Other_GSSs: CIT-HSP-2341D11.TR
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.
FEATURES
    source
    1. .733
        Location/Qualifiers
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /clone="2341D11"
            /sex="Male"
            /cell_type="Sperm"
            /clone_lib="CIT-HSP"
            /note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
            HindIII"

Query Match 3.9%; Score 41.4; DB 9; Length 733;
Best Local Similarity 53.4%; Pred. No. 6.1;

ORIGIN
```

```
Matches 87; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
QY 129 GCAGTAGCAAAATGGAACATTTGGTAAAAACAATACTGATGACCAAAATTCAGGTGACAAAT 188
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 469 GCAGCAGGAAATGAAATGATAGATTAAATTTAAAGAGAGAGAAAAATATCAATAGTAGTTT 410
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 189 GCTACTGAATTAGTTTCAGAGCATTTCAATAGGGAATAATGCAACAACCTCATATAAAGTT 248
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 409 TAAATAAATATGTTTCAGAAATCTGAAGGGTAAAAAATGAATAATCAATAAAGGAA 350
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 249 CTAGATGGAAGAAATTCACATTAATAGATGCAATGCTAGGAG 291
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 349 CAAGAAATTTATGCACTAAAGCAGATTAAATGAAATAATAAAG 307
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 10
LOCUS CNS0006Y 1101 bp DNA linear GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC:
BACR34L01 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL073010
VERSION AL073010.1 GI:4952872
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Direct Submission
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL Bp 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammeter in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw ap, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
    source
    1. .1101
        Location/Qualifiers
            /organism="Drosophila melanogaster"
            /mol_type="genomic DNA"
            /db_xref="taxon:7227"
            /clone="BACR34L01"
            /clone_lib="RPCI-98"
            /note="end : T7"

Query Match 3.9%; Score 41; DB 10; Length 1101;
Best Local Similarity 43.1%; Pred. No. 8.9;
Matches 119; Conservative 13; Mismatches 144; Indels 0; Gaps 0;
QY 79 AAACCCCAACGAGTGGAAACAACACAGCCACATTATGCTGGGACACCATGCGTAGTACCA 138
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 302 AAACCCCAACGAGTGGAAACAACACAGCCACATTATGCTGGGACACCATGCGTAGTACCA 361
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 139 ATGGAACATTTGGTAAAAACAATAACTGATGACCAAAATTCAGGTGACAAATGCTACTGAAT 198
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 362 TACAAACAANWACACCAANWATATTTATATATAATAATAATAATAATAATAATAATAATA 421
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 199 TAGTTCAGACATTTTCAATAGGGAAATAATGCAACAACCTCATATAAAGTTCTAGATGGAA 258
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

```

Db      422 CAATACCAATATATATATATATCCYATATTCACATACACATATATATATACAA 481
QY      259 GAAATTCACATTAATAGATGCAATGCTAGGAGACCCCACTGTGTCTTCCAGTATG 318
Db      482 CMAATTRGGCTKTTTAAACTCTTTTCSAATASGCGCCCACTKTSAGAATTKCAAACTC 541
QY      319 AGAATTCGGACCTCTTCATAGAAGAAGACGCGCTT 354
Db      542 AGACTTGGAGAACAAATAGSAGCAASATCAT 577

```

```

RESULT 11
AG429915/c
LOCUS   Mus musculus molossinus DNA, clone:MSMg01-304L12.TJ, genomic survey
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1
Abe, K., Noguchi, H., Tagawa, K., Yuzuriha, M., Toyoda, A., Kojima, T.,
Ezawa, K., Saitou, N., Hattori, M., Sakaki, Y., Moriwaki, K. and
Shiroishi, T.
Contribution of Asian mouse subspecies Mus musculus molossinus to
genomic constitution of strain C57BL/6J, as defined by BAC-end
sequence-SNP analysis
Genome Res. 14 (12), 2439-2447 (2004)
15574823
2 (bases 1 to 1362)
Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
Direct Submission
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
[E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170]
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center.
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY
Vector      : pBACe3.6
R.Site 1    : ECoRI
R.Site 2    : ECoRI.
FEATURES
source
1..1362
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-304L12.TJ"
/sex="male"
/tissue_types="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"

```

```

ORIGIN
Query Match      3.9%; Score 41; DB 10; Length 1362;
Best Local Similarity 34.2%; Pred.No. 9.5;
Matches 179; Conservative 0; Mismatches 345; Indels 0; Gaps 0;
QY      536 AACAAATCTGGAACCTTTACCCCACTTGAATGTGACAAATGCTTAACAATAAAATTT 595

```

```

Db      1086 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1027
QY      596 CGCAAACTATACATCTGGGGGATTCTATCACCAGCTCAACCAACAGCAGACAGATT 655
Db      1026 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 967
QY      656 GTACATCCAGAATACAGGACGAGTAAACAGTCTCAACAAAAAGAGTCAACAAACGATAGT 715
Db      966 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 907
QY      716 CCCTAATATCGGATCTAGACCGTGGTTCAGGCGTCAATCAGGCGAGGATAGCATATATCTG 775
Db      906 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 847
QY      776 GACCATTTGTAACCTGGAGATATCTTAATGATAACAGTAATGCGCACTTAGTTGCACC 835
Db      846 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 787
QY      836 CGCGGGATATTTAAATTTGAAACACAGGAAAGCTCTGTATGAGATCAGATGCACCCAT 895
Db      786 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 727
QY      896 AGCATTGTGTCTCGAATGTATTACCAAAATGGAAGCATCCCAACGACAAACCAT 955
Db      726 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 667
QY      956 TCAAAATGTCAACAAAGTTACATATGGAATATGCCCACTATATATCAGGCAAAACACTTT 1015
Db      666 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 607
QY      1016 AAAGCTGGCCACTGGGATGAGGAATATACCAAGAAAGCAATCA 1059
Db      606 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 563

```

```

RESULT 12
CX894805/c
LOCUS   JGI CAAM6042.rev NIH XGC tropTe3 Xenopus tropicalis cDNA clone
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Xenopus tropicalis (western clawed frog)
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 750)
Richardson, P., Lucas, S., Rokhsar, D., Dettter, J.C., Ng, D.C.,
Brokstein, P. and Lindquist, E.A.
DOE Joint Genome Institute Xenopus tropicalis EST project
Unpublished (2004)
On Feb 4, 2005 this sequence version replaced gi:58634149.
Other ESTs: JGI CAAM6042.fwd
Contact: Lindquist, E.A., Richardson, P.
DOE Joint Genome Institute
2800 Mitchell Drive, Walnut Creek, CA 94598, USA
Tel: 925 296 5600
Fax: 925 296 5710
Email: cdna@jgi-psf.org
Tissue Procurement: Timothy Grammer (Richard M. Harland Laboratory,
University of California, Berkeley:
http://tropicalis.berkeley.edu/home)
cDNA Library Preparation: DOE Joint Genome Institute:
http://www.jgi.doe.gov
DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
Clone Distribution: I.M.A.G.E. Consortium/LLNL:
http://image.llnl.gov
Naming Conventions: EST name is generated by the concatenation of
the JGI Clone Id and the direction of sequencing. The suffix '.rev'
indicates a reverse sequencing read of the insert. It does not
necessarily reflect the orientation of the insert.

```


THIS PAGE BLANK (USPTO)